photonics

Review

The Role of Auxiliary Stages in Gaussian Quantum Metrology

Danilo Triggiani [1], Paolo Facchi [2,3] and Vincenzo Tamma [1,4,*]

[1] School of Mathematics and Physics, University of Portsmouth, Portsmouth PO1 3QL, UK; danilo.triggiani@port.ac.uk
[2] Dipartimento di Fisica and MECENAS, Università di Bari, I-70126 Bari, Italy; paolo.facchi@ba.infn.it
[3] INFN, Sezione di Bari, I-70126 Bari, Italy
[4] Institute of Cosmology and Gravitation, University of Portsmouth, Portsmouth PO1 3FX, UK
* Correspondence: vincenzo.tamma@port.ac.uk

Abstract: The optimization of the passive and linear networks employed in quantum metrology, the field that studies and devises quantum estimation strategies to overcome the levels of precision achievable via classical means, appears to be an essential step in certain metrological protocols achieving the ultimate Heisenberg-scaling sensitivity. This optimization is generally performed by adding degrees of freedom by means of auxiliary stages, to optimize the probe before or after the interferometric evolution, and the choice of these stages ultimately determines the possibility to achieve a quantum enhancement. In this work we review the role of the auxiliary stages and of the extra degrees of freedom in estimation schemes, achieving the ultimate Heisenberg limit, which employ a squeezed-vacuum state and homodyne detection. We see that, after the optimization for the quantum enhancement has been performed, the extra degrees of freedom have a minor impact on the precision achieved by the setup, which remains essentially unaffected for networks with a larger number of channels. These degrees of freedom can thus be employed to manipulate how the information about the structure of the network is encoded into the probe, allowing us to perform quantum-enhanced estimations of linear and non-linear functions of independent parameters.

Keywords: quantum metrology; quantum sensing; distributed parameter; heisenberg limit; typicality; gaussian metrology; squeezing; estimation of functions

check for updates

Citation: Triggiani, D.; Facchi, P.; Tamma, V. The Role of Auxiliary Stages in Gaussian Quantum Metrology. *Photonics* **2022**, *9*, 345. https://doi.org/10.3390/photonics9050345

Received: 30 March 2022
Accepted: 6 May 2022
Published: 14 May 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

In recent years, much attention has been put in the study of metrological schemes that exploit quantum resources, such as entanglement and squeezing, to enhance the sensitivity in the estimation of physical properties beyond the possibilities of classical strategies, with applications to imaging [1,2], thermometry [3,4], mapping of magnetic fields [5,6] and gravitational waves detection [7], among others. One of the most emblematic quantum enhancements sought in quantum metrology is the renown Heisenberg limit [8–28], which consists in achieving a scaling of the estimation error in the number N of probes (typically photons, or atoms) of order of $1/N$, which surpasses the classical (or shot-noise) limit $1/\sqrt{N}$.

Gaussian metrology, which specializes in the study of estimation schemes employing Gaussian states of light and squeezing as metrological resource [29–32], represents a promising path towards a feasible quantum-enhancement in estimation strategies and the Heisenberg-scaling sensitivity [33–41]. It exploits the possibility to reduce the intrinsic noise of the electromagnetic field quadratures below the quantum fluctuations of the vacuum. Such a reduced noise, together with relatively easy-to-implement experimental procedures to produce these squeezed-noise states, and their increased robustness to decoherence compared to entangled states, make the Gaussian approach of great interest for short-term applications of quantum technologies. A particular case analysed by quantum metrology is the estimation of a single unknown parameter that appears within a given optical linear network multiple times, affecting for example different interferometric components [34–36,38,42–48] (see Figure 1). This is the case of unknown temperatures or magnitudes of the electromagnetic field, which modifies the physical properties of the optical parts composing the network within the regime of passive and linear evolution of the probe. The field investigating this type of schemes is generally referred to as distributed metrology, since the unknown parameter is effectively distributed among multiple components of the network.

On the other hand, estimation schemes based on Gaussian metrology usually incur in the challenge of adaptivity, i.e., the fact that the protocol depends on the value of the parameter that tries to estimate [44,49–52]. A typical approach to deal with adaptivity in Gaussian metrology consists in limiting the values that the unknown parameter can take, for example restricting the working range of the estimation scheme only to small values of the parameter, a condition that is common in a typical interferometric setup [43,46,48,53]. However, this solution unfortunately excludes certain experimental situations which require the ability to perform a quantum-enhanced estimation of the unknown parameter without imposing restrictions on its value.

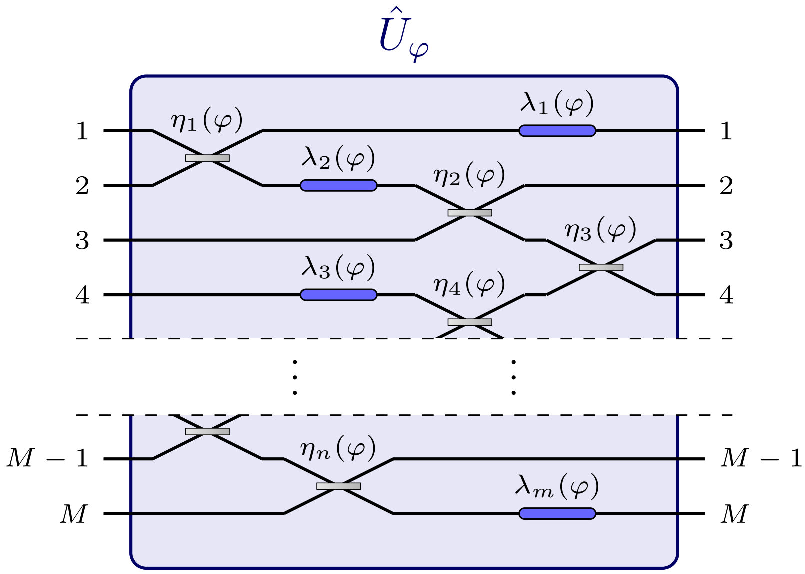

Figure 1. Example of a passive and linear network $\hat{U}_\varphi$ which depends on a single global parameter φ [35]. The parameter can be thought as a physical property of an external agent (e.g., temperature, electromagnetic field) which affects multiple components, possibly of different nature, of the network [35,38].

Remarkably, it has been found in a recent work that is possible to achieve Heisenberg-scaling sensitivity in the estimation of a given unknown parameter distributed in an arbitrary M-channel network with a Gaussian scheme that only requires a classical knowledge of the parameter to optimize the network, i.e., the unknown parameter must be known with a prior precision that can be achieved with a classical estimation strategy [35]. To implement this scheme, a single auxiliary optical network is required in order to correctly refocus the probe, a squeezed-vacuum state, into the only output port observed through homodyne detection after the interferometric evolution, and a classical knowledge on the unknown parameter is required to engineer this auxiliary network. In other words, since in general an arbitrary M-channel linear network which encodes an unknown parameter does not refocus the probe into a single output port, some degrees of freedom must be introduced in the network by adding an auxiliary stage, which need to be optimized through a classical estimation strategy, which assures that the refocusing is correctly performed [35].

Within this scheme, provided that the optimization of the degrees of freedom has been performed, and the probe is correctly refocused in the only observed output port, it is always possible to add a second auxiliary network, which represents further extra degrees of freedom introduced in the optical network. One may wonder how the choice of this further auxiliary stage, and thus the presence of extra degrees of freedom, can influence the estimation scheme. Remarkably, in Ref. [36] it has been shown that the extra degrees of freedom introduced with a second auxiliary network do not have a major impact on the precision of the estimation of the unknown parameter. In particular, for networks with a large number of channels, it was shown that a random choice of the non-optimized auxiliary network leaves the precision of the estimation scheme essentially unaltered [36]. Therefore, this auxiliary stage and the extra degrees of freedom can be used to manipulate the way the information on the structure of the network is encoded in the probe.

In Ref. [37] it has been shown that it is possible to employ the previous scheme to achieve the Heisenberg-scaling sensitivity in the estimation of suitable functions of multiple parameters, and to exploit the degrees of freedom introduced by the non-optimized network to control the form of the function that can be estimated. Despite, in principle, complications may arise due to the presence of multiple parameters, representing multiple sources of uncertainty, which require a more complex mathematical formalism to describe a multi-parameter scenario, it is possible to employ the further degrees of freedom to manipulate the way the information on the structure of the network is encoded into the probe, and ultimately control the function of the parameters that is estimated at the Heisenberg-scaling sensitivity [37]. It is worth mentioning that, although the task of estimating functions of unknown parameters can be easily performed estimating separately each single parameter, and then evaluating the function in the data post-processing analysis, the ability to directly estimate a global property (e.g., spatial average of a field, field-gradients, or non-linear functions) allows one to avoid the waste of resources to obtain superfluous information on each single parameter, hence its relevance in applications, such as evaluation of averages of magnetic fields or temperatures.

This review is organized as follows. First, we briefly review the general scheme presented in Ref. [35] for the estimation of a single distributed parameter in Section 2. In Section 3 we discuss the effect of the presence of a second auxiliary network on the precision of the estimation, showing that the extra degrees of freedom introduced after the optimization of the refocusing has been performed do not essentially affect the precision of the estimation scheme [36]. Lastly, we will see in Section 4 how it is possible to exploit the exceeding degrees of freedom of the auxiliary stages which are not employed to optimize the network in a network with multiple parameters to change the function that can be estimated at the Heisenberg-scaling precision [37]. We conclude presenting two examples. The first is a 2-channel network which allows to estimate a function of three parameters (two optical phases and a beam-splitter reflectivity) parametrized by some quantities that can be chosen arbitrarily through the auxiliary stages. We will see that, according to the choice of the auxiliary networks, the function estimated can be linear or non-linear in the three parameters. The second is a scheme for the estimation of any linear combination of parameters with positive weights. In particular, we will show how it is possible to employ this scheme when the unknown parameter are not only phase-shifts, but also reflectivities of beam-splitters, or more in general phases acquired through complex local networks.

2. Distributed-Parameter Quantum-Enhanced Estimation

We will start introducing the Gaussian estimation scheme for M-channel networks which achieves the Heisenberg-scaling sensitivity recently proposed in Ref. [35]. Let us consider an arbitrary linear passive network $\hat{U}_\varphi$ which depends on a single parameter φ possibly distributed among several components of the network. The preparation of the input probe consists in the injection of a single-mode squeezed vacuum state in the first port of an auxiliary linear and passive network $\hat{V}_{\text{in}}$, which is used to scatter the photons injected among all the modes. The input state in our protocol is therefore given by

$$|\psi_0\rangle = \hat{V}_{\text{in}}\hat{S}_1(r)|\text{vac}\rangle \tag{1}$$

where $\hat{S}_1(r) = \mathrm{e}^{\frac{r}{2}(\hat{a}_1^2 - \hat{a}_1^{\dagger 2})}$ is the squeezing operator associated with the first channel with squeezing parameter $r > 0$, and $|\mathrm{vac}\rangle$ is the M-channel vacuum state. The average number N of photons injected in the network is thus $N = \sinh^2 r$. At the output of $\hat{U}_\varphi$, a further auxiliary network $\hat{V}_{\mathrm{out}}$ is employed to refocus all the photons into a single mode, namely the first one, in order to capture all the information about the parameter in a single channel, at which homodyne detection is performed with a given local oscillator phase θ. For a linear passive unitary $\hat{U}$, it is possible to introduce the $M \times M$ unitary matrix U whose elements U_{ij} represent the single-photon transition amplitudes from the i-th input port to the j-th output port, defined by the map

$$\hat{U}^\dagger \hat{a}_i^\dagger \hat{U} = \sum_{j=1}^{M} U_{ij} \hat{a}_j^\dagger. \tag{2}$$

We now introduce the $M \times M$ unitary matrix $u_\varphi = V_{\mathrm{out}} U_\varphi V_{\mathrm{in}}$ associated with the evolution through the whole network $\hat{V}_{\mathrm{out}} \hat{U}_\varphi \hat{V}_{\mathrm{in}}$. Then, we can write the probability

$$P_\varphi = \left|(u_\varphi)_{11}\right|^2 = \left|(V_{\mathrm{out}} U_\varphi V_{\mathrm{in}})_{11}\right|^2 \tag{3}$$

that a photon injected in the first port of $\hat{V}_{\mathrm{in}}$ comes out from the first port of $\hat{V}_{\mathrm{out}}$, and the phase

$$\gamma_\varphi = \arg[(u_\varphi)_{11}] = \arg[(V_{\mathrm{out}} U_\varphi V_{\mathrm{in}})_{11}] \tag{4}$$

accumulated through this interferometric evolution. One can show, by employing Cramér-Rao analysis [54,55], that it is possible to achieve Heisenberg-scaling sensitivity in the estimation of φ if the conditions

$$\gamma_\varphi - \theta = \pm\frac{\pi}{2} + \frac{k}{N} + \mathcal{O}\left(\frac{1}{N^2}\right), \tag{5a}$$

$$P_\varphi = 1 - \frac{\ell}{N} + \mathcal{O}\left(\frac{1}{N^2}\right), \tag{5b}$$

are satisfied, where $\ell \geqslant 0$ and $k \neq 0$ are arbitrary but both independent of N, and where θ is an optimal choice for the local oscillator phase. Under conditions (5), the ultimate precision achievable by this scheme with any estimator $\tilde{\varphi}$ after ν iterations of the measurement is given by

$$\mathrm{Var}[\tilde{\varphi}] \geqslant \frac{1}{\nu \mathcal{F}(\varphi)} = \frac{1}{8\nu \varrho(k,\ell)(\partial_\varphi \gamma_\varphi)^2 N^2}, \tag{6}$$

where $\mathcal{F}(\varphi)$ is the Fisher information associated with this estimation scheme, and

$$\varrho(k,\ell) = \left(\frac{8k}{1 + 16k^2 + 4\ell}\right)^2, \tag{7}$$

is an N-independent factor which reaches its maximum $\varrho = 1$ at $k = 1/4$ and $\ell = 0$.

One can show that it is possible to optimize the refocusing network $\hat{V}_{\mathrm{out}}$ with only classical prior information on the parameter φ, namely after a classical strategy is employed to perform a prior coarse estimation of the unknown parameter.

In the following we will focus on the relation between the non-optimized auxiliary network $\hat{V}_{\mathrm{in}}$, which yields further degrees of freedom in the linear network of the scheme, and the precision shown in Equation (6). Then, we will show how it is possible to employ these extra degrees of freedom to manipulate the function of multiple parameters that can be estimated with this scheme.

3. Typicality of Quantum Enhanced Sensitivity

In the previous section we have presented a generic protocol that allows us to achieve the Heisenberg limit in the estimation of a parameter φ distributed in an arbitrary network $\hat{U}_\varphi$

when conditions (5) are met. In particular, in order to satisfy condition (5b), either the scattering stage $\hat{V}_{\mathrm{in}}$ or the refocusing stage $\hat{V}_{\mathrm{out}}$ needs to be optimized after a classical prior estimation of the parameter φ is carried out. The remaining auxiliary stage is thus left completely arbitrary, and one may wonder how the choice of this stage can influence the precision of the estimation, and in particular the pre-factor $(\partial_\varphi \gamma_\varphi)^2$ appearing in the sensitivity in Equation (6). More precisely, it may happen that a particularly unfortunate choice of the non-optimized stage causes this pre-factor to vanish, for example if this auxiliary stage transforms the optical mode of the probe into a mode which is insensible on or independent of the value of the parameter (A trivial example is the case where the unitary matrix describing the network is $U_\varphi = \mathrm{diag}(1, \exp(\mathrm{i}\varphi))$, and the auxiliary stage is the identity $V_{\mathrm{in}} = \mathbb{1}_2$, for $M = 2$. In this case the probe is left in the first channel of the network, which does not depend on φ.) In this section we will see that, for an arbitrary choice of the non-optimized auxiliary stage, the pre-factor $(\partial_\varphi \gamma_\varphi)^2$ tends to be far from zero, meaning that small values are mostly unlikely, especially for networks with a large number M of channels [36]. For simplicity, we will explicitly consider the case in which the refocusing stage $\hat{V}_{\mathrm{out}}$ has been optimized to satisfy condition (5b) while $\hat{V}_{\mathrm{in}}$ is left arbitrary, but similar considerations can be done in the opposite scenario due to the symmetry of the problem.

3.1. The Role of the Generator G_φ

We can link the pre-factor $(\partial_\varphi \gamma_\varphi)^2$, appearing in Equation (6), to the derivative of the matrix element $(u_\varphi)_{11}$

$$\begin{aligned} \left|(\partial_\varphi u_\varphi)_{11}\right|^2 &= \left|\left(\partial_\varphi \sqrt{P_\varphi} + i(\partial_\varphi \gamma_\varphi)\sqrt{P_\varphi}\right)\mathrm{e}^{\mathrm{i}\gamma_\varphi}\right|^2 \\ &= (\partial_\varphi \sqrt{P_\varphi})^2 + (\partial_\varphi \gamma_\varphi)^2 P_\varphi. \end{aligned} \tag{8}$$

When condition (5b) holds, Equation (8) simplifies to

$$(\partial_\varphi \gamma_\varphi)^2 = \left|(\partial_\varphi u_\varphi)_{11}\right|^2 + \mathcal{O}\left(\frac{1}{N}\right), \tag{9}$$

so that the two quantities are equal up to a term of order $1/N$. Condition (5b) can be recast in terms of a constraint on the form of V_{out}

$$(V_{\mathrm{out}})_{1i} = (V_{\mathrm{in}}^\dagger U_\varphi^\dagger)_{1i} + \mathcal{O}\left(\frac{1}{\sqrt{N}}\right) \tag{10}$$

If we now introduce the (generally φ-dependent) generator

$$G_\varphi := \mathrm{i} U_\varphi^\dagger \frac{\partial U_\varphi}{\partial \varphi} \tag{11}$$

of the unitary matrix U_φ, we can further manipulate the pre-factor $(\partial_\varphi \gamma_\varphi)^2$. Employing the definition of G_φ in Equation (11), and the relation in Equation (10), we can write

$$\begin{aligned} \left|(\partial_\varphi u_\varphi)_{11}\right|^2 &= \left|\left(V_{\mathrm{out}} \frac{\partial U_\varphi}{\partial \varphi} V_{\mathrm{in}}\right)_{11}\right|^2 \\ &= \left|(V_{\mathrm{out}} U_\varphi G_\varphi V_{\mathrm{in}})_{11}\right|^2 \\ &= (V_{\mathrm{in}}^\dagger G_\varphi V_{\mathrm{in}})_{11}^2 + \mathcal{O}\left(\frac{1}{N}\right) \end{aligned} \tag{12}$$

Equations (9) and (12) conveniently express the pre-factor $(\partial_\varphi \gamma_\varphi)^2$ as a function of the generator G_φ of the network and a unitary matrix $U = V_{\text{in}}$ independent of the optimized stage V_{out}, so that ultimately we can write for large N

$$(\partial_\varphi \gamma_\varphi)^2 \sim f(U, G_\varphi) = (U^\dagger G_\varphi U)_{11}^2 \tag{13}$$

and thus the Fisher information appearing in the sensitivity in Equation (6) becomes

$$\mathcal{F}(\varphi) \sim 8\varrho(k, \ell) f(U, G_\varphi) N^2. \tag{14}$$

It is easy to see that the pre-factor $f(U, G_\varphi)$ can be written as the square of a convex combination of the eigenvalues $\{g_i\}_{i=1,\ldots,M}$ of the generator G_φ. In fact, if we call V_φ the matrix whose columns are the eigenvectors of G_φ, so that $D \equiv \text{diag}(g_1, \ldots, g_M) = V_\varphi^\dagger G_\varphi V_\varphi$, we can recast the pre-factor in terms of the eigenvalues $\{g_i\}_{i=1,\ldots,M}$

$$(U^\dagger G_\varphi U)_{11}^2 = \left(\sum_{j=1}^{M} w_j g_j \right)^2, \tag{15}$$

where $w_j = \left| (V_\varphi^\dagger U)_{j1} \right|^2$, with $\sum_j w_j = 1$ by the unitarity of $V_\varphi^\dagger U$. If we suppose, without lack of generality, that the eigenvalues g_i are ordered so that $|g_i| \geqslant |g_{i+1}|$, the maximum value of the pre-factor is achieved for

$$w_1 \equiv \left| (V_\varphi^\dagger U)_{11} \right|^2 = 1, \tag{16}$$

namely when the first column of U coincides (up to a complex phase) with the first eigenvector of G_φ, belonging to the eigenvalue with the larger absolute value. Recalling that U represents the action of the input auxiliary stage (see Equation (12)), and that the elements of its first column coincides with the transition amplitudes from the first input channel (see Equation (2)), we can understand the meaning of the condition (16): the input stage must be chosen in order to maximize the effect of the network $\hat{U}_\varphi$ on the probe, which must evolve under the optical mode (not necessarily coinciding with a physical channel) which is most sensitive to the variations of φ—i.e., corresponding to the eigenvalue of G_φ with maximum absolute value. For a choice of U satisfying condition (16) (e.g., for $U = V_\varphi$), the pre-factor in Equation (13) coincides with the highest eigenvalue of the generator squared

$$\max_U f(U, G_\varphi) \equiv f_{\max} = g_1^2 \equiv \left\| G_\varphi \right\|^2, \tag{17}$$

namely the squared norm of the generator G_φ. This coincides with the pre-factor of the maximum Quantum Fisher information for Gaussian states found in Ref. [44], meaning that the scheme presented in Section 2 is the optimal Gaussian strategy when also the auxiliary stage $\hat{V}_{\text{in}} \equiv U$ is optimized according to Equation (16).

3.2. Typical Behaviour of the Pre-Factor in the Heisenberg Scaling

Although the condition for the optimal choice of the unitary U has been easily found, the eigenvectors of G_φ in general depend on the value of the distributed parameter φ, and on the structure of the network $\hat{U}_\varphi$, and so do the solutions of Equation (16). On the other hand, one may be interested in the case in which no prior knowledge on the structure of $\hat{U}_\varphi$—and thus on G_φ— is given. In general, with these assumptions, finding the optimal stage that maximizes the pre-factor $f(U, G_\varphi)$ and satisfies condition (16) is impossible. Instead, in this scenario, it becomes more appropriate to study the behaviour of the pre-factor for arbitrary choices of the auxiliary network, and ultimately of the non-fixed degrees of freedoms introduced with it. In particular, one may be interested in knowing how likely the value of $f(U, G_\varphi)$ is equal or close to zero, when the auxiliary stage is chosen at random. However, to introduce concepts such as 'how likely' and 'at random', we must somehow endow the set of all the auxiliary linear network with a probability measure. Conveniently, there is already a mathematical structure which well represents this set, which we have already extensively employed: every M-channel linear and passive network is described through the action of a $M \times M$ unitary matrix, and the 'composition rules' of linear networks is well represented by the composition rules of $\mathrm{U}(M)$, the group of $M \times M$ unitary matrices.

Since we are supposing that we do not possess any prior information on the structure of U_φ, all the auxiliary stages U are equivalent candidate to achieve the optimal value for the pre-factor. For this reason, we can endow $\mathrm{U}(M)$ with its Haar measure $\mathcal{P}$ [56]. This is the measure on the space of the $M \times M$ unitary matrices that generalizes the uniform distribution on finite intervals of $\mathbb{R}$. In fact, the Haar measure on $\mathrm{U}(M)$ is defined so that the measure $\mathcal{P}$ of any open subset $\mathcal{U} \subset \mathrm{U}(M)$ is invariant under left or right unitary transformations, namely $\mathcal{P}(\mathcal{U}) = \mathcal{P}(U'\mathcal{U}) = \mathcal{P}(\mathcal{U}U')$, for every $U' \in \mathrm{U}(M)$. In other words, it is a measure that only depends on the 'size' of the subsets of $\mathrm{U}(M)$. Once we have chosen the prescription to sample the network U, we are able to evaluate statistical properties of the pre-factor $f(U, G_\varphi)$. In particular, we show in Appendix A that

$$\mathbb{E}_{\mathcal{P}}[f(U, G_\varphi)] = \frac{\mathrm{Tr}\left(G_\varphi^2\right) + \mathrm{Tr}\left(G_\varphi\right)^2}{M(M+1)} \tag{18}$$

is the expectation value of the pre-factor (13) over random choices of the matrix U, with respect to the Haar measure. In the case of a generator proportional to the identity $G_\varphi = \|G_\varphi\| \mathbb{1}_M$, which corresponds to a network made of M identical copies of single-channel unitaries acting in parallel (This condition is satisfied for example for the metrological scheme introduced in Ref. [12], although in such a case the generator G_φ is independent of φ), we have $\mathrm{Tr}\left(G_\varphi\right)^2 = M^2\|G_\varphi\|^2$ and $\mathrm{Tr}\left(G_\varphi^2\right) = M\|G_\varphi\|^2$, so that the average value of the pre-factor equals the maximum $f_{\max} = \|G_\varphi\|^2$ in Equation (17). Indeed, with this kind of network, the auxiliary stage which distribute the probe on all M channels becomes irrelevant, since the network acts identically in each channel. Indeed, in such a case any unitary matrix $V_\varphi \equiv U$ diagonalizes the generator G_φ and therefore Equation (16) is satisfied for any $V_{\mathrm{in}} \equiv U$.

In general, an average value of the pre-factor close to the maximum in Equation (17) is favourable to a smaller value, since it implies a better Fisher information and a better precision in average. The only case for which $\mathbb{E}_{\mathcal{P}}[f(U, G_\varphi)]$ is equal to zero is when the whole generator is vanishing, occurrence happening only for networks U_φ which do not actually depend on the unknown parameter. A generator G_φ with small eigenvalues—i.e., a network which is not very sensitive to the variations of the parameter φ— would cause the average in Equation (18) to decrease, diminishing the average precision of the estimation scheme. However, it is possible to find a lower bound on the average value (18) using Jensen's inequality $\mathbb{E}[X^2] \geqslant \mathbb{E}[X]^2$ to obtain

$$\mathbb{E}_{\mathcal{P}}[f(U, G_\varphi)] \geqslant \mathbb{E}_{\mathcal{P}}[(U^\dagger G_\varphi U)_{11}]^2 = \left[\frac{\mathrm{Tr}(G_\varphi)}{M}\right]^2 \tag{19}$$

(see Appendix A). We notice how the right-hand side term in Equation (19) is the squared average of the eigenvalues of G_φ. This means that if we are able to control the value of the average of the generator, say in such a way that it is larger of a certain fraction $R > 0$ of the norm of the generator, i.e., $\mathrm{Tr}\left[G_\varphi\right]/M \geqslant R\|G_\varphi\|$, then it follows from Equation (19) that we can assure that the average of the pre-factor is larger than a fraction R^2 of its maximum value $\mathbb{E}_{\mathcal{P}}[f(U, G_\varphi)] \geqslant R^2\|G_\varphi\|^2 \equiv R^2 f_{\max}$.

Even though we may be able to control the average of $f(U, G_\varphi)$ through Equation (19), it may happen that the *typical* values that the pre-factor takes are far from $\mathbb{E}_{\mathcal{P}}[f(U, G_\varphi)]$, for random choices of the unitary U. A paradigmatic example of a random variable which typically takes very different values than its average is a quantity which can only be equal to 0 or 1 with equal probabilities: in such case, its average 1/2 is never an effective outcome of the random variable, let alone typical. Fortunately, this is not the case for the pre-factor $f(U, G_\varphi)$, which is instead a well-behaved function of the random unitary U. In fact, we see in Appendix A that $f(U, G_\varphi)$ is instead typical for networks with many channels, i.e., that it is possible to apply results on the concentration of measure in high-dimensional spaces, which assure that $f(U, G_\varphi)$ becomes almost constant for random choices of $U \in \mathrm{U}(M)$ for large M, and thus it concentrates around its average value

$$\mathcal{P}(|f - \mathbb{E}_{\mathcal{P}}[f]| \geqslant \varepsilon) \leq 2\exp\left(-\frac{AM}{\|G_\varphi\|^4}\varepsilon^2\right), \qquad \forall\varepsilon > 0 \tag{20}$$

with $A = (72\pi^3)^{-1}$. This results tell us that, for an arbitrary choice of auxiliary stage $U = V_{\mathrm{in}}$, the value of the pre-factor $(\partial_\varphi \gamma_\varphi)^2 \equiv f(U, G_\varphi)$ appearing in the Fisher information in Equation (14) is with overwhelming probability close to its average, for networks with a large enough number of channels M (see Figure 2). Moreover, Equation (19) shows that it is possible to bound from below the average $\mathbb{E}_{\mathcal{P}}(f(U, G_\varphi))$ of the pre-factor, if some control on the average of the eigenvalues of G_φ is possible. This shows that, beside very unlikely exceptions, the choice of the non-optimized stage is mostly irrelevant as for the precision of the estimation scheme. In the next section, we will see how it is possible to exploit the additional degrees of freedom introduced by the non-optimized network to manipulate how the information on the structure of the network is encoded in the probe. This will give us some freedom in choosing the function of a given set of parameters that can be estimated with the Heisenberg limit precision.

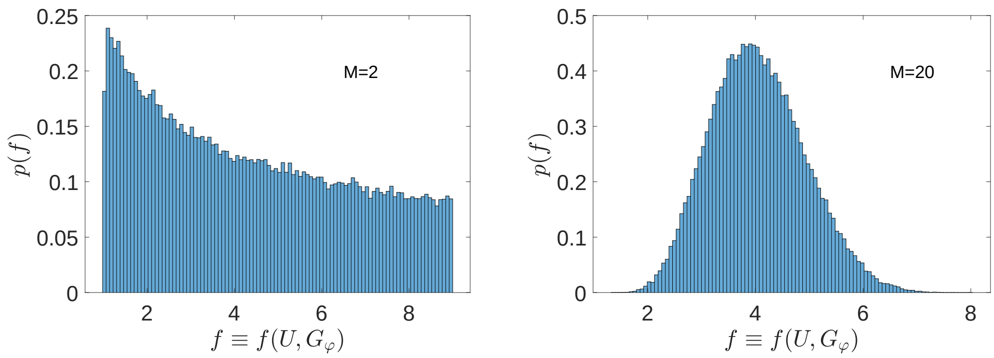

Figure 2. Histograms of the pre-factor $f(U, G_\varphi)$ in Equation (13) for $M = 2$ (**left**) and $M = 20$ (**right**), obtained numerically with 10^5 samplings of U with respect of the unitary Haar measure [36]. The generator G_φ chosen is a diagonal matrix with half 1s and half 3s as entries. The histograms are normalized to the unity. We can see that in the histogram in the right, the values of the pre-factor are more concentrated around its average $\mathbb{E}_{\mathcal{P}}[f(U, G_\varphi)] \simeq 4$.

4. Estimation of Functions of Parameters

In the previous section, we have seen that it is always possible to add a φ-independent auxiliary stage to the estimation setup shown in Section 2, which introduces degrees of freedom that essentially do not affect the precision of the estimation scheme, especially for networks with a large number of channels. A natural question that may arise is whether the same setup can be employed in achieving the Heisenberg limit when the number of unknown parameters affecting the arbitrary network increases, and if it is possible to employ these degrees of freedom to select a specific function of the parameters that can be measured with Heisenberg-scaling sensitivity. Although such task can be performed estimating separately each single parameter, and then evaluating the function during the data analysis, the ability to directly estimate a global property (e.g., spatial average of a field, field-gradients, or non-linear functions) allows us to not waste resources to obtain superfluous information on each single parameter. However, it cannot be excluded in principle that complications may arise due to the presence of multiple sources of uncertainty, complications which already materialize starting from the more complex mathematical formalism required to describe the multi-parameter scenario [57,58].

In this section we will describe a scheme for the estimation of functions of multiple parameters encoded in a generic linear passive network [37]. We will show that, employing a single squeezed vacuum state and a single homodyne detector, i.e., the same probe and measurement of the setup shown in Section 2, it is possible to reach the Heisenberg limit in the estimation of functions of the parameters, satisfying conditions that are similar to the ones found for the single-parameter scheme. We find also in this scenario that a classical knowledge on the unknown parameters is required to optimize the network through the use of auxiliary stages, allowing us to conceive two-steps protocols achieving the Heisenberg limit. Moreover, we will see how the exceeding degrees of freedom of the auxiliary stages which are not employed to optimize the network can be used to change the functional dependence between the quantity that can be estimated and the unknown parameters.

Once we will have described the estimation scheme and discussed the conditions that need to be met in order to reach the Heisenberg limit, we will present two examples. The first is a 2-channel network which allows us to estimate a function of three parameters, of which two optical phases and a beam-splitter reflectivity, parametrized by some quantities that can be chosen arbitrarily through the auxiliary stages employed. We will see that, according to the choice of the auxiliary networks, the function estimated can be linear or non-linear in the three parameters. The second is a scheme for the estimation of any linear combination of parameters with positive weights. In particular, we will show how it is possible to employ this scheme when the unknown parameter are not only phase-shifts, but also reflectivities of beam-splitters, or more in general phases acquired through complex local networks.

4.1. Setup

Let us consider a M-channel linear and passive network $\hat{U}_{\boldsymbol{\varphi}}$ which depends on p unknown parameters $\boldsymbol{\varphi} = (\varphi_1, \ldots, \varphi_p)$. The parameters $\boldsymbol{\varphi}$ may represent certain physical properties associated with each component of the network, such as reflectivities of beam-splitters, or phase-shift magnitudes, or they may be the values of external non-uniform fields which influence several components of the network, such as the temperature and the electromagnetic field (see Figure 3). The action of the network can be described with the usual unitary matrix representation $U_{\boldsymbol{\varphi}}$, defined through Equation (2). The probe employed is the same as shown in Equation (1), injected in a single input port of an auxiliary network $\hat{V}_{\text{in}}$, say the first, with $N = \sinh^2 r$ number of photons in average in the probe. In order to infer some information on the parameters from the interferometric evolution of the probe, we will perform homodyne measurements at a single output port, say the first, of the quadrature $\hat{x}_\theta$, where θ is the phase of the local oscillator. Similarly to the setup described in Section 2, we will consider in this model the presence of a further refocusing network $\hat{V}_{\text{out}}$ acting on the probe after the evolution given by the network $\hat{U}_{\boldsymbol{\varphi}}$ respectively. Intuitively, the role of the stage $\hat{V}_{\text{in}}$ is to distribute the probe among all the channels of the network, while $\hat{V}_{\text{out}}$ refocuses the probe into the only output port which is observed. However, in light of the results of typicality presented in Section 3, which showed that the overall precision of the single-parameter estimation scheme is essentially not affected by the choice of the non-optimized network for a large number M of channels, we will exploit the remaining degrees of freedom in the two auxiliary stages to manipulate how the information about the structure of the network $\hat{U}_{\boldsymbol{\varphi}}$ and on the parameters $\boldsymbol{\varphi}$ is encoded into the probe.

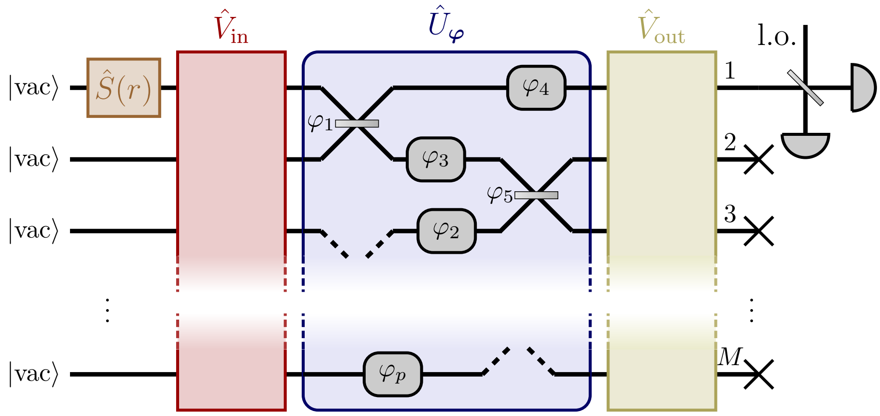

Figure 3. Diagram of the setup described in Section 4.1 [37]. A squeezed vacuum state is injected in the first input port of a network composed of a first auxiliary stage $\hat{V}_{\text{in}}$, a linear and passive network $\hat{U}_{\boldsymbol{\varphi}}$ which depends on multiple unknown parameters $\boldsymbol{\varphi}$, and a second auxiliary stage $\hat{V}_{\text{out}}$. The two auxiliary stages are linear and passive networks whose purpose is to manipulate how the information on $\boldsymbol{\varphi}$ and on the structure of the network is encoded into the probe, and to refocus it in the only output channel observed through homodyne detection. This setup reaches the Heisenberg limit in the estimation of the overall phase acquired by the probe, which is a function of the parameters $\boldsymbol{\varphi}$ that can be manipulated through the choice of $\hat{V}_{\text{in}}$ and $\hat{V}_{\text{out}}$.

Since the photons of the probe are all injected in the first channel, and only the first output port is observed, the only relevant element of the unitary matrix $u_{\boldsymbol{\varphi}} = V_{\text{out}} U_{\boldsymbol{\varphi}} V_{\text{in}}$— representing the action of the whole setup on the probe— is $(u_{\boldsymbol{\varphi}})_{11}$, namely the transition amplitude from the first input to the first output port. We will employ the parametrization

$$(u_{\boldsymbol{\varphi}})_{11} \equiv (V_{\text{out}} U_{\boldsymbol{\varphi}} V_{\text{in}})_{11} = \sqrt{P_{\boldsymbol{\varphi}}} \mathrm{e}^{\mathrm{i} f(\boldsymbol{\varphi})}, \tag{21}$$

which emphasizes the two relevant physical quantities, i.e., the transition probability $P_{\boldsymbol{\varphi}} := \left|(u_{\boldsymbol{\varphi}})_{11}\right|^2$ and the phase $f(\boldsymbol{\varphi}) := \arg(u_{\boldsymbol{\varphi}})_{11}$ acquired by the probe through the network, which is in general a function of the unknown parameters $\boldsymbol{\varphi}$. We will see later that the function $f(\boldsymbol{\varphi})$ can be estimated at the Heisenberg limit. After the interferometric evolution, the squeezed variance hence becomes $\hat{x}_{f(\boldsymbol{\varphi})+\pi/2}$. If the quadrature $\hat{x}_\theta$ is observed through homodyne detection, the probability distribution $p_{\boldsymbol{\varphi}}(x)$ which governs the outcomes of the measurement is Gaussian, due to the Gaussian nature of the probe [30], and centred in zero, due to the absence of displacement. We thus write the Gaussian probability distribution

$$p_{\boldsymbol{\varphi}}(x) = \frac{1}{\sqrt{2\pi\sigma_{\boldsymbol{\varphi}}^2}} \mathrm{e}^{-\frac{x^2}{2\sigma_{\boldsymbol{\varphi}}^2}}, \tag{22}$$

with variance $\sigma_{\boldsymbol{\varphi}}$ calculated in Appendix B

$$\sigma_{\boldsymbol{\varphi}}^2 = \frac{1 - P_{\boldsymbol{\varphi}}}{2} + \frac{P_{\boldsymbol{\varphi}}}{2} [\cosh(2r) + \cos(2f(\boldsymbol{\varphi}) - 2\theta) \sinh(2r)], \tag{23}$$

and can be thought as the average between the noises of the vacuum and of the squeezed states, weighted by the factors $1 - P_{\boldsymbol{\varphi}}$ and $P_{\boldsymbol{\varphi}}$ respectively.

The presence of multiple independent parameters imposes the multi-parameter approach for the analysis of the ultimate precisions achievable with this setup, and the use of the Fisher information matrix [54,55],

$$\mathcal{F}(\boldsymbol{\varphi})_{ij} = \int \mathrm{d}x \, p_{\boldsymbol{\varphi}}(x) \left(\frac{\partial}{\partial \varphi_i} \ln p_{\boldsymbol{\varphi}}(x) \right) \left(\frac{\partial}{\partial \varphi_j} \ln p_{\boldsymbol{\varphi}}(x) \right) \tag{24}$$

By plugging Equation (22) into Equation (24) one gets

$$\mathcal{F}(\boldsymbol{\varphi}) = \frac{1}{2\sigma_{\boldsymbol{\varphi}}^4}\left(\nabla\sigma_{\boldsymbol{\varphi}}^2\right)\left(\nabla\sigma_{\boldsymbol{\varphi}}^2\right)^{\mathrm{T}}, \tag{25}$$

where $\nabla = \nabla_{\boldsymbol{\varphi}} = (\frac{\partial}{\partial\varphi_1}, \ldots, \frac{\partial}{\partial\varphi_p})^{\mathrm{T}}$. Having the probability distribution in Equation (22) a null displacement, all the information on the parameters $\boldsymbol{\varphi}$ is encoded in the variance $\sigma_{\boldsymbol{\varphi}}$ through the transition probability $P_{\boldsymbol{\varphi}}$ and the phase acquired $f(\boldsymbol{\varphi})$. It is thus convenient to separate the derivative of the variance in Equation (23) in two contributions

$$\nabla\sigma_{\boldsymbol{\varphi}}^2 = (\partial_P\sigma_{\boldsymbol{\varphi}}^2)\nabla P_{\boldsymbol{\varphi}} + (\partial_f\sigma_{\boldsymbol{\varphi}}^2)\nabla f(\boldsymbol{\varphi}), \tag{26}$$

where ∂_P and ∂_f denote the partial derivatives with respect to P and f, so that we easily obtain from Equation (23)

$$\partial_P\sigma_{\boldsymbol{\varphi}}^2 = \frac{1}{2}(-1 + \cosh(2r) + \cos(2f(\boldsymbol{\varphi}) - 2\theta)\sinh(2r)), \tag{27a}$$

$$\partial_f\sigma_{\boldsymbol{\varphi}}^2 = -P_{\boldsymbol{\varphi}}\sin(2f(\boldsymbol{\varphi}) - 2\theta)\sinh(2r). \tag{27b}$$

4.2. Heisenberg Scaling

We can see from Equations (23)–(27) that, if no specific conditions are imposed on the network $u_{\boldsymbol{\varphi}}$, the Fisher information matrix $\mathcal{F}(\boldsymbol{\varphi})$ cannot generally reach the Heisenberg limit. This is due to the presence of the squared variance $\sigma_{\boldsymbol{\varphi}}^4$ at the denominator of Equation (25), which may grow with order N^2, with N number of photons injected, since

$$\begin{aligned}\sigma_{\boldsymbol{\varphi}}^2 &= \frac{1 - P_{\boldsymbol{\varphi}}}{2} + \frac{P_{\boldsymbol{\varphi}}}{2}\left[1 + 2N + 2\cos(2f(\boldsymbol{\varphi}) - 2\theta)\sqrt{N(N+1)}\right]\\ &= NP_{\boldsymbol{\varphi}}(1 + \cos(2f(\boldsymbol{\varphi}) - 2\theta)) + \mathcal{O}(1),\end{aligned} \tag{28}$$

while the derivatives in Equations (27) only contain terms of the type $\sinh(2r)$ and $\cosh(2r)$, which are of order N. Thus it arises the need to impose conditions which prevent the variance of the observed quadrature to grow with the number of photons. We show in Appendix C that this setup reaches the Heisenberg limit in the estimation of $f(\boldsymbol{\varphi})$ if similar conditions to the ones in Equation (5) for the single-parameter scenario are met. In particular, the conditions are

$$f(\boldsymbol{\varphi}) - \theta = \pm\frac{\pi}{2} + \frac{k}{N} + \mathcal{O}\left(\frac{1}{N^2}\right), \qquad k \neq 0, \tag{29a}$$

$$P_{\boldsymbol{\varphi}} = 1 - \frac{\ell}{N} + \mathcal{O}\left(\frac{1}{N^2}\right), \qquad \ell \geq 0, \tag{29b}$$

with k and ℓ arbitrary factors which are independent of N. In Section 4.3, we will discuss about the meaning of the conditions in Equation (29), exploring their consequences and highlighting the similarities with the single-parameter protocol in Section 2. In Appendix C we show that, when conditions (29) are met, the Fisher information matrix in Equation (25) becomes

$$\mathcal{F}(\boldsymbol{\varphi}) \sim 8\varrho(k,\ell)N^2(\nabla f(\boldsymbol{\varphi}))(\nabla f(\boldsymbol{\varphi}))^{\mathrm{T}}, \tag{30}$$

where $\varrho(k,\ell)$ is the same N-independent constant factor defined in Equation (7) for the single-parameter estimation.

Despite the presence of the factor N^2 in the Fisher information matrix in Equation (30), it is generally impossible to reach the Heisenberg limit in the estimation of each of the p parameters $\boldsymbol{\varphi}$. In fact, we can easily see that the (asymptotic) expression of the Fisher information matrix (30) is non-invertible: the (column) vector $\nabla f(\boldsymbol{\varphi})$ is trivially the only eigenvector associated with a non-vanishing eigenvalue of the Fisher information matrix.

As discussed in Refs. [59,60], a singular Fisher information matrix is symptomatic of the presence of parameters which do not admit estimators with finite variances, and the traditional multi-parameter Cramér-Rao bound [54,55]

$$\mathrm{Cov}(\tilde{\varphi}_i, \tilde{\varphi}_j) \geqslant \frac{1}{\nu}\mathcal{F}(\boldsymbol{\varphi})^{-1}_{ij} \tag{31}$$

is not applicable due to the non-invertibility of $\mathcal{F}(\boldsymbol{\varphi})$. In fact, only specific functions $\psi(\boldsymbol{\varphi})$ of the parameters $\boldsymbol{\varphi}$ admit unbiased estimators with finite variance [59,60]. In particular, specialized to the Fisher information matrix of the problem at hand in Equation (30), the functions $\psi(\boldsymbol{\varphi})$ which admit finite variance are the ones satisfying [59]

$$\nabla\psi(\boldsymbol{\varphi}) \propto \nabla f(\boldsymbol{\varphi}). \tag{32}$$

The solutions of Equation (32) for all the possible values of $\boldsymbol{\varphi}$ are of the form $\psi(\boldsymbol{\varphi}) \equiv a_1 f(\boldsymbol{\varphi}) + a_2$, with $a_{1/2}$ independent of $\boldsymbol{\varphi}$, and in particular $\psi(\boldsymbol{\varphi}) \equiv f(\boldsymbol{\varphi})$ belongs to this family, which entails the fact that the function $f(\boldsymbol{\varphi})$ can be estimated with finite variance. We can finally evaluate the Cramér-Rao bound associated with any estimator $\tilde{f}$ of $f(\boldsymbol{\varphi})$, obtaining in the asymptotic regime (see Appendix E)

$$\mathrm{Var}_{\boldsymbol{\varphi}}[\tilde{f}] \geqslant \frac{1}{\nu}\frac{1}{8\varrho(k,\ell)N^2}, \tag{33}$$

which reaches the Heisenberg limit in the mean number of photons N.

Lastly, it is worth to mention that we can saturate the Cramér-Rao bound in the limit of large samples, namely it is always possible to find an estimator, the maximum-likelihood estimator $\tilde{f}_{\mathrm{MLE}}$, which is unbiased and efficient in the asymptotic regime of large samples $\nu \to +\infty$ [55]. In order to find the maximum-likelihood estimator, we need to maximize the Likelihood function

$$\mathcal{L}(\boldsymbol{\varphi}; \boldsymbol{x}) = \prod_{i=1}^{\nu} p_{\boldsymbol{\varphi}}(x_i) = \frac{1}{(2\pi\sigma^2_{\boldsymbol{\varphi}})^{\nu/2}} \exp\left(-\frac{|\boldsymbol{x}|^2}{2\sigma^2_{\boldsymbol{\varphi}}}\right) \tag{34}$$

In Appendix D we see that the solution which maximizes the Likelihood function (34) is given by the estimator $\tilde{f}_{\mathrm{MLE}}$ satisfying

$$\sigma^2(\tilde{f}_{\mathrm{MLE}}) = S^2(\boldsymbol{x}), \tag{35}$$

where $\sigma^2(f)$ is the variance in Equation (23) as a function of $f \equiv f(\boldsymbol{\varphi})$, supposing that $P_{\boldsymbol{\varphi}}$ satisfies Equation (29b) and is known, while $S(\boldsymbol{x})^2$ is the usual sample variance

$$S^2(\boldsymbol{x}) = \frac{1}{\nu}\sum_{i=1}^{\nu} x_i^2. \tag{36}$$

Inverting the function $\sigma^2(f)$, we can obtain the explicit expression of the maximum-likelihood estimator in this regime

$$\tilde{f}_{\mathrm{MLE}}(\boldsymbol{x}) = \theta + \frac{1}{2}\left(2n\pi \pm \arccos\left(\frac{(2S(\boldsymbol{x})^2 - 1) - 2P_{\boldsymbol{\varphi}}\sinh^2 r}{2P_{\boldsymbol{\varphi}}\sinh r\cosh r}\right)\right), \tag{37}$$

with n integer. We notice that the presence of the cosine function, which is invertible only on intervals of its argument of the type $[n\pi, (n+1)\pi[$, requests a prior knowledge on the argument $2f(\boldsymbol{\varphi}) - 2\theta$ in order to choose the correct value of n in Equation (37). However, we will discuss in the next section that a classical coarse estimation of the parameters $\boldsymbol{\varphi}$ is required in order to optimize the network and satisfy condition (29b). In other words, the error $\delta\varphi$ committed in the prior coarse estimation must be of order $1/\sqrt{N}$, decreasing at the SQL with the number of photons N. For a large enough N, $\delta\varphi$ will be small enough to unequivocally choose the correct interval of invertibility of the cosine, and thus the correct n in Equation (37).

4.3. On the Conditions for the Heisenberg Scaling

Despite the presence of multiple parameters, the substantial similarity of conditions (29) with the single-parameter counterparts in Equation (5) allows us to draw the same consideration discussed for the single-parameter scheme in Ref. [35]. Condition (29a) is a minimum resolution requirement on the tuning of local-oscillator phase θ, which must be controlled with steps of order $1/N$. Moreover, the requirement $k \neq 0$ in Equation (29a) implies that θ must be tuned in such a way to measure a quadrature field $\hat{x}_\theta$ which is slightly different from the minimum-variance quadrature $\hat{x}_{f(\boldsymbol{\varphi})\pm\pi/2}$. This can be explained by the fact that the minimum variance—i.e., $\sigma^2_{\boldsymbol{\varphi}}$ for $\theta = f(\boldsymbol{\varphi}) \pm \pi/2$— is a stationary point for variations of $f(\boldsymbol{\varphi})$, and hence its gradient in Equation (26) is vanishing for $P_{\boldsymbol{\varphi}}$ close to its maximum and $\sigma^2_{\boldsymbol{\varphi}}$ to its minimum.

Condition (29b) is a requirement on the refocusing of the probe into the only observed channel. In order for the variance in Equation (23) of the observed quadrature to be 'squeezed'—i.e., $\sigma^2_{\boldsymbol{\varphi}} \sim 1/N$— the contribution of the vacuum $(1-P_{\boldsymbol{\varphi}})/2$ must be of order $1/N$. Moreover, similar considerations also for the prior knowledge on the parameter required for the optimization of the refocusing network can be drawn. First of all, condition (29b) can be satisfied by only optimizing a single auxiliary stage, while choosing arbitrarily the other. We can call $\boldsymbol{\varphi}_{\text{cl}}$ the result of a prior coarse estimation of the parameters $\boldsymbol{\varphi}$ required to optimize a single auxiliary stage, say $\hat{V}_{\text{out}} \equiv \hat{V}_{\text{out}}(\boldsymbol{\varphi}_{\text{cl}})$. The single-photon probability transition $P_{\boldsymbol{\varphi}}$ can be written as the squared modulus of the scalar product of the two vectors $U_{\boldsymbol{\varphi}} V_{\text{in}} \boldsymbol{e}_1$ and $V_{\text{out}}(\boldsymbol{\varphi}_{\text{cl}})^\dagger \boldsymbol{e}_1$, with $\boldsymbol{e}_1 = (1, 0, \ldots, 0)^{\text{T}}$

$$P_{\boldsymbol{\varphi}} = \left| \boldsymbol{e}_1^{\text{T}} V_{\text{out}}(\boldsymbol{\varphi}_{\text{cl}}) U_{\boldsymbol{\varphi}} V_{\text{in}} \boldsymbol{e}_1 \right|^2 \equiv \eta(\boldsymbol{\varphi}, \boldsymbol{\varphi}_{\text{cl}}), \tag{38}$$

where $\eta(\boldsymbol{\varphi}, \boldsymbol{\varphi}_{\text{cl}})$ is a smooth function of $\boldsymbol{\varphi}$ and $\boldsymbol{\varphi}_{\text{cl}}$ which is maximized for $\boldsymbol{\varphi}_{\text{cl}} = \boldsymbol{\varphi}$, since $\hat{V}_{\text{out}}(\boldsymbol{\varphi}_{\text{cl}} = \boldsymbol{\varphi})$ would satisfy $\eta(\boldsymbol{\varphi}, \boldsymbol{\varphi}) \equiv 1$. For small deviations $\delta\boldsymbol{\varphi} = \boldsymbol{\varphi} - \boldsymbol{\varphi}_{\text{cl}}$ of the coarse estimation from the true values of the parameters $\boldsymbol{\varphi}$, we can expand Equation (38)

$$\eta(\boldsymbol{\varphi}, \boldsymbol{\varphi} - \delta\boldsymbol{\varphi}) = 1 - \underbrace{\sum_{i=1}^{p} \frac{\partial \eta(\boldsymbol{\varphi}, \boldsymbol{x})}{\partial x_i}\bigg|_{\boldsymbol{x}=\boldsymbol{\varphi}}}_{0} \delta\varphi_i + \frac{1}{2} \sum_{i,j=1}^{p} \frac{\partial^2 \eta(\boldsymbol{\varphi}, \boldsymbol{x})}{\partial x_i \partial x_j}\bigg|_{\boldsymbol{x}=\boldsymbol{\varphi}} \delta\varphi_i \delta\varphi_j + (\mathcal{O}(\delta\boldsymbol{\varphi}))^3, \tag{39}$$

where the gradient $\nabla_{\boldsymbol{x}} \eta(\boldsymbol{\varphi}, \boldsymbol{x})$ is zero for $\boldsymbol{x} = \boldsymbol{\varphi} = \boldsymbol{\varphi}_{\text{cl}}$. Thus, also in the presence of multiple unknown parameter $\boldsymbol{\varphi}$, if the errors $\delta\boldsymbol{\varphi}$ in the prior estimations $\boldsymbol{\varphi}_{\text{cl}}$ are of order $1/\sqrt{N}$, it becomes possible to engineer the refocusing stage $\hat{V}_{\text{out}}(\boldsymbol{\varphi}_{\text{cl}})$ that satisfies condition (29b), and thus that allows to reach the Heisenberg limit, similarly to the single-parameter estimation protocol.

4.4. Examples of Quantum-Enhanced Estimation of Functions

The presence of two auxiliary stages $\hat{V}_{\text{in}}$ and $\hat{V}_{\text{out}}$, and the need to optimize only one of them in order to satisfy condition (29b)—and ultimately to reach the Heisenberg limit— entail the possibility to exploit the remaining degrees of freedom in the network to manipulate how the information on the parameters $\boldsymbol{\varphi}$ are encoded in the probe. In this section, two examples are proposed, which make use of these degrees of freedom to allow us to choose the function to be estimated from a family of functions of some parameters $\boldsymbol{\varphi}$. The first example is a 2-channel network for the estimation at the Heisenberg limit of non-linear functions of three parameters, of which two are magnitudes of phase-shifts and one is the reflectivity of a beam-splitter. The second is a network for the estimation of linear combination with positive weights of an arbitrary number of parameters.

4.4.1. Non-Linear Functions

We here consider a 2-channel network $\hat{U}_{\boldsymbol{\varphi}}$ for the estimation of non-linear functions $f(\boldsymbol{\varphi};\boldsymbol{\alpha})$ of the reflectivity of a beam-splitter φ_1 and the magnitudes of two phase-shifts $\varphi_{2/3}$, with reference to Figure 4. The functions $f(\boldsymbol{\varphi};\boldsymbol{\alpha})$ are parametrized by the quantities $\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \alpha_3, \alpha_4)$, which can be implemented arbitrarily, with the only condition that $\alpha_1 - \alpha_2 = \alpha_4 - \alpha_3 \equiv \Delta\alpha$. The possibility to choose the quantities $\boldsymbol{\alpha}$ stems from the presence of abundant degrees of freedom in the auxiliary networks $\hat{V}_{\text{in}}$ and $\hat{V}_{\text{out}}$, which are not employed for the optimization of the network needed to satisfy Equation (29b). The protocol employed is the same considered in Section 4.1, with a single squeezed-vacuum state with $N = \sinh^2 r$ average number of photons injected in the first input port of the overall network, with r squeezing parameter of the probe, and only the first output channel observed through homodyne detection, to measure the quadrature $\hat{x}_\theta$, where θ is the phase of the homodyne local oscillator.

We can write the matrices representing the action of the beam-splitter and the phase-shifts in the network $\hat{U}_{\boldsymbol{\varphi}}$ as

$$
\begin{aligned}
U_{\text{BS}}(\varphi_1) &= \exp(\mathrm{i}\varphi_1\sigma_2) = \begin{pmatrix} \cos\varphi_1 & \sin\varphi_1 \\ -\sin\varphi_1 & \cos\varphi_1 \end{pmatrix} \\
U_{\text{PS}}(\varphi_2,\varphi_3) &= \exp\left(\mathrm{i}\frac{\varphi_2+\varphi_3}{2}\mathbb{1}_2 + \mathrm{i}\frac{\varphi_2-\varphi_3}{2}\sigma_3\right) = \begin{pmatrix} \exp(\mathrm{i}\varphi_2) & 0 \\ 0 & \exp(\mathrm{i}\varphi_3) \end{pmatrix}
\end{aligned} \tag{40}
$$

respectively, where σ_i, $i = 1,2,3$, is the i-th Pauli matrix and $\mathbb{1}_2$ is the 2×2 identity matrix, so that the network $\hat{U}_{\boldsymbol{\varphi}}$ is represented by the matrix

$$
U_{\boldsymbol{\varphi}} = U_{\text{PS}}(\varphi_2,\varphi_3)U_{\text{BS}}(\varphi_1) = \begin{pmatrix} \cos\varphi_1\exp(\mathrm{i}\varphi_2) & \sin\varphi_1\exp(\mathrm{i}\varphi_2) \\ -\sin\varphi_1\exp(\mathrm{i}\varphi_3) & \cos\varphi_1\exp(\mathrm{i}\varphi_3) \end{pmatrix}. \tag{41}
$$

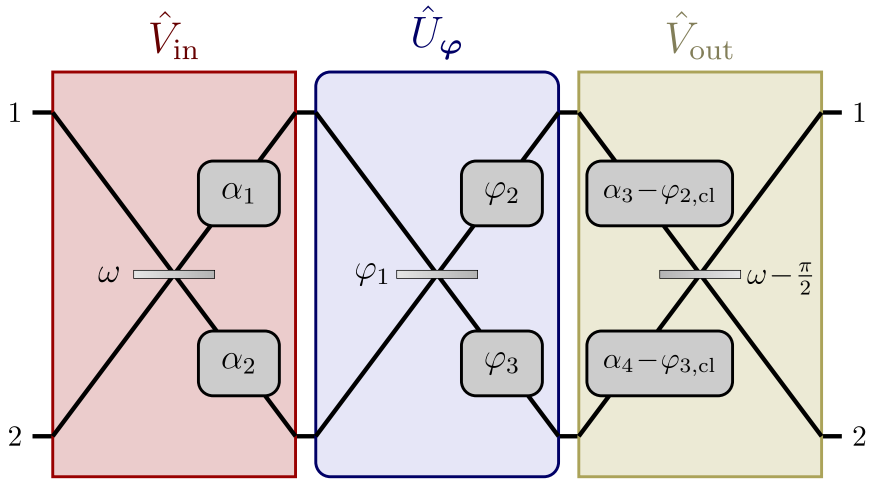

Figure 4. A 2-channel example of a network which allows to estimate at the Heisenberg limit certain functions of the reflectivity φ_1 of a beam-splitter and the magnitudes $\varphi_{2/3}$ of phase-shifts [37]. The form of the function can be manipulated by changing the values of the arbitrary parameters $\boldsymbol{\alpha}$, with $\alpha_1 - \alpha_2 = \alpha_4 - \alpha_3 = \Delta\alpha$ (see Equation (46)), while the value of ω is shown in Equation (43). Similarly to the setup for the single-parameter model described in Section 2, a classical knowledge $\boldsymbol{\varphi}_{\text{cl}}$ of the unknown parameters suffices to optimize the network.

The quantity $\Delta\alpha = \alpha_1 - \alpha_2$ that we will use to parametrize the family of functions $f(\boldsymbol{\varphi};\boldsymbol{\alpha})$ is the relative phase between the arms of the input auxiliary stage $\hat{V}_{\text{in}} \equiv \hat{V}_{\text{in}}(\alpha_1,\alpha_2)$, which is overall described by the unitary matrix

$$
V_{\text{in}}(\alpha_1,\alpha_2) = U_{\text{PS}}(\alpha_1,\alpha_2)U_{\text{BS}}(\omega). \tag{42}
$$

This auxiliary network consists of two phase-shifts of arbitrary magnitudes α_1 and α_2 in the first and second channel, and a beam splitter with reflectivity

$$\omega = \frac{1}{2}\arctan\left(\frac{\cos(\varphi_{\mathrm{cl},1})}{\sin(\varphi_{\mathrm{cl},1})\cos\Delta\alpha}\right), \tag{43}$$

which can be engineered once a classical estimation $\varphi_{\mathrm{cl},1}$ of the unknown reflectivity φ_1 of the beam-splitter in $\hat{U}_{\boldsymbol{\varphi}}$ has been carried out, namely after a coarse prior estimation so that the error $\delta\varphi_1 = \varphi_1 - \varphi_{\mathrm{cl},1}$ is of order $1/\sqrt{N}$. The refocusing auxiliary network $\hat{V}_{\mathrm{out}}$ is described, with reference to Figure 4, by the unitary matrix

$$V_{\mathrm{out}}(\alpha_3,\alpha_4) = U_{\mathrm{BS}}(\omega-\pi/2)U_{\mathrm{PS}}(\alpha_3-\varphi_{\mathrm{cl},2},\alpha_4-\varphi_{\mathrm{cl},3}), \tag{44}$$

where the quantity $\Delta\alpha = \alpha_4 - \alpha_3$ enters in the relative phase of the two channels, this time changed in sign. We can see from the expression of V_{out} that this auxiliary stage in general depends on the classical estimation $\boldsymbol{\varphi}_{\mathrm{cl}}$ of the unknown parameters, namely on the results of a coarse prior estimation for which the committed errors $\delta\boldsymbol{\varphi} = \boldsymbol{\varphi} - \boldsymbol{\varphi}_{\mathrm{cl}}$ are of order $1/\sqrt{N}$.

We will now see that the Heisenberg limit can be achieved in the estimation of the complex phase $f(\boldsymbol{\varphi};\boldsymbol{\alpha}) \equiv \arg(V_{\mathrm{out}}(\alpha_3,\alpha_4)U_{\boldsymbol{\varphi}}V_{\mathrm{in}}(\alpha_1,\alpha_2))$ of the network $\hat{u}_{\boldsymbol{\varphi}}$ depicted in Figure 4, showing that it satisfies condition (29b). In fact, employing Equations (41)-(44), we can explicitly evaluate the transition amplitude

$$\begin{aligned}(u_{\boldsymbol{\varphi}})_{11} = \mathrm{e}^{\mathrm{i}\left(\alpha_1+\alpha_3+\frac{\delta\varphi_2+\delta\varphi_3}{2}\right)}\Bigg(&\cos\left(\frac{\delta\varphi_2-\delta\varphi_3}{2}\right)\sin 2\omega\cos\varphi_1\\ &+\cos\left(\Delta\alpha-\frac{\delta\varphi_2-\delta\varphi_3}{2}\right)\cos 2\omega\sin\varphi_1 + \mathrm{i}\sin\left(\Delta\alpha-\frac{\delta\varphi_2-\delta\varphi_3}{2}\right)\sin\varphi_1\Bigg),\end{aligned} \tag{45}$$

where the quantity ω is defined in Equation (43). In Appendix F we see that the transition probability $P_{\boldsymbol{\varphi}} \equiv |u_{\boldsymbol{\varphi}}|^2_{11}$ actually satisfies the condition (29b) on the network for the Heisenberg limit, which means that the auxiliary networks correctly operate on the probe so that it gets refocused on the only observed channel. In Appendix F the complex phase is evaluated

$$f(\boldsymbol{\varphi};\boldsymbol{\alpha}) = \arctan\left(\frac{\mathrm{Im}(u_{\boldsymbol{\varphi}})_{11}}{\mathrm{Re}(u_{\boldsymbol{\varphi}})_{11}}\right) = \alpha_1+\alpha_3+\frac{\delta\varphi_2+\delta\varphi_3}{2} + \arctan(\Phi), \tag{46}$$

with

$$\Phi = \frac{\sin\varphi_1\sin\left(\Delta\alpha-\frac{\delta\varphi_2-\delta\varphi_3}{2}\right)\sqrt{1-\sin^2(\varphi_{\mathrm{cl},1})\sin^2\Delta\alpha}}{\cos\varphi_1\cos\varphi_{\mathrm{cl},1}\cos\left(\frac{\delta\varphi_2-\delta\varphi_3}{2}\right)+\sin\varphi_1\sin\varphi_{\mathrm{cl},1}\cos\Delta\alpha\cos\left(\Delta\alpha-\frac{\delta\varphi_2-\delta\varphi_3}{2}\right)}, \tag{47}$$

which is in general a non-linear function of the parameters $\boldsymbol{\varphi}$. We see from Equations (46) and (47) how the choice of the arbitrary values of the parameters $\boldsymbol{\alpha}$ affects the functional dependence of the phase $\arg(u_{\boldsymbol{\varphi}})_{11}$ on the parameters $\boldsymbol{\varphi}$, and thus of the quantity which can be estimated at the Heisenberg limit. In particular, beside the term $\alpha_1+\alpha_3$ which simply adds an overall phase, the relative phase $\Delta\alpha$ affects the functional dependence between the quantity Φ in Equation (47) and $\boldsymbol{\varphi}$.

For example, for $\Delta\alpha = \pi/2$ the beam-splitters in the auxiliary stages $\hat{V}_{\mathrm{in}}$ and $\hat{V}_{\mathrm{out}}$ become balanced—i.e., $\omega = \pm\pi/4$ from Equation (43)—while the function of $\boldsymbol{\varphi}$ in Equation (46) we can estimate reduces to

$$f(\boldsymbol{\varphi};\Delta\alpha=\pi/2) = \alpha_1+\alpha_2-\frac{\varphi_{\mathrm{cl},2}+\varphi_{\mathrm{cl},3}}{2}+\varphi_1+\frac{\varphi_2+\varphi_3}{2}, \tag{48}$$

thus becoming a linear function of the parameters. From Equation (45), we can evaluate the transition probability $P_{\boldsymbol{\varphi}} = |(u_{\boldsymbol{\varphi}})|^2_{11}$ for the choice of $\Delta\alpha = \pi/2$, and for $\omega = \pi/4$

$$P_{\boldsymbol{\varphi}} = \cos^2\left(\frac{\delta\varphi_2 - \delta\varphi_3}{2}\right) = 1 - \frac{(\delta\varphi_2 - \delta\varphi_3)^2}{4} + \mathcal{O}(\delta\boldsymbol{\varphi})^3, \tag{49}$$

which satisfies condition (29b) as expected, since the error in the coarse estimation are assumed to be classical, i.e., so that $\delta\boldsymbol{\varphi} = \mathcal{O}(1/\sqrt{N})$. Comparing Equations (29b) and (49), and assuming that $\delta\varphi_i = k_i/\sqrt{N}$ with k_i independent of N, it is easy to evaluate the factor

$$\ell = \frac{(k_2 - k_3)^2}{4} \tag{50}$$

which enters through $\varrho(k, \ell)$ in Equation (7) in the Cramér-Rao bound in Equation (33). Moreover we notice that, if the two phase-shifts are completely known quantities, so that we are able to perfectly balance them with the auxiliary stage $\hat{V}_{\text{out}}$ with $\delta\varphi_{2/3} = 0$, the overall network $\hat{u}_{\boldsymbol{\varphi}}$ in Figure 4 reduces (up to a global phase $(\alpha_1 + \alpha_2)/2$) to a setup which transform the reflectivity φ_1 in an optical delay, without any prior information on the parameter. In fact in this case, the overall phase $f(\boldsymbol{\varphi}; \Delta\alpha = \pi/2)$ in Equation (48) becomes $\alpha_1 + \alpha_2 + \varphi_1$. In Section 4.4.2 we will make use of this type of networks specifically for this purpose, and be able to treat the reflectivities of beam-splitter as if they were simple phase-shift.

For $\Delta\alpha = 0$, the reflectivity ω in Equation (43) becomes $\omega = -\varphi_{\text{cl},1}/2 - \pi/4$, while the function estimated at the Heisenberg limit in Equation (48) reads

$$\begin{aligned} f(\boldsymbol{\varphi}; \Delta\alpha = 0) &= \alpha_1 + \alpha_2 + \frac{\delta\varphi_2 + \delta\varphi_3}{2} + \arctan\left(-\tan\left(\frac{\delta\varphi_2 + \delta\varphi_3}{2}\right)\frac{\sin\varphi_1}{\cos\delta\varphi_1}\right) \\ &= \alpha_1 + \alpha_2 + \frac{\delta\varphi_2 + \delta\varphi_3}{2} - \frac{\delta\varphi_2 + \delta\varphi_3}{2}\sin\varphi_1 + \mathcal{O}(\delta\boldsymbol{\varphi})^3, \end{aligned} \tag{51}$$

where the term $\mathcal{O}(\delta\boldsymbol{\varphi})^3$ can be neglected since it is of order $N^{-3/2}$ or smaller, i.e., beyond the Heisenberg limit resolution. In particular, we notice from Equation (51) that $f(\boldsymbol{\varphi}; \Delta\alpha = 0)$ is a relatively simple non-linear function, in which we can find the products $\varphi_2 \sin\varphi_1$ and $\varphi_3 \sin\varphi_1$, between each phase-shifts and the transmittivity amplitude of the beam-splitter. If we once again evaluate the transition probability $P_{\boldsymbol{\varphi}} = |(u_{\boldsymbol{\varphi}})|^2_{11}$ from Equation (45) for $\Delta\alpha = 0$ we obtain

$$P_{\boldsymbol{\varphi}} = 1 - \delta\varphi_1^2 - \frac{(\delta\varphi_2 - \delta\varphi_3)^2\cos^2\varphi_1}{4} + \mathcal{O}(\delta\boldsymbol{\varphi})^3, \tag{52}$$

which satisfies condition (29b) as expected, since $\delta\boldsymbol{\varphi}$ is of order $1/\sqrt{N}$. In particular, we can easily evaluate the N-independent factor

$$\ell = k_1^2 + \frac{(k_2 - k_3)^2\cos^2\varphi_1}{4} \tag{53}$$

which enters in the Cramér-Rao bound in Equation (33), if we assume $\delta\varphi_i = k_i/\sqrt{N}$ with k_i independent of N. If moreover $\varphi_1 = 0$, the network in Figure 4 reduces to a balanced Mach-Zehnder, which allows us to estimate the average of the two phase-shifts since the overall phase in Equation (51) reduces, for $\varphi_1 = 0$, to

$$f(\boldsymbol{\varphi}; \Delta\alpha = 0) = \alpha_1 + \alpha_2 - \frac{\varphi_{\text{cl},2} - \varphi_{\text{cl},3}}{2} + \frac{\varphi_2 + \varphi_3}{2}, \tag{54}$$

where the average of the two phase-shifts φ_2 and φ_3 sums a known quantity, which can thus be subtracted during the estimation without affecting the overall precision. In the next example, we will generalize such scheme, considering a similar network, but with an arbitrary number of channels and generally unbalanced beam-splitters.

4.4.2. Linear Combinations of Arbitrary Parameters

We now consider a network $\hat{U}_{\boldsymbol{\varphi}}$ which depends on M independent parameters $\boldsymbol{\varphi}$ (see Figure 5), that allows us to estimate with a precision at the Heisenberg limit any linear combination

$$L(\boldsymbol{\varphi}) = \sum_{i=1}^{M} w_i \varphi_i \equiv \boldsymbol{w} \cdot \boldsymbol{\varphi} \tag{55}$$

with positive weights $\boldsymbol{w} = (w_1, \dots, w_M)$. As also discussed previously in this chapter, the ability to change arbitrarily the weights stems from the presence of degrees of freedom in the auxiliary stages $\hat{V}_{\text{in}}$ and $\hat{V}_{\text{out}}$ that are not employed to refocus the probe in the only observed channel, i.e., to satisfy condition (29b). At first, we will suppose that the parameters $\boldsymbol{\varphi}$ can either be magnitudes of phase-shifts or reflectivities of beam-splitters. Later, we will show how it is possible to generalize the scheme for an arbitrary set of parameters. We will also suppose, without loss of generality, that the weights $\boldsymbol{w}$ sum to one, namely that the linear combination in Equation (55) is a convex sum. To estimate a generic linear combination, it would suffice to rescale the estimated convex sum, causing a rescaling of the error in the estimation which does not affect the Heisenberg limit.

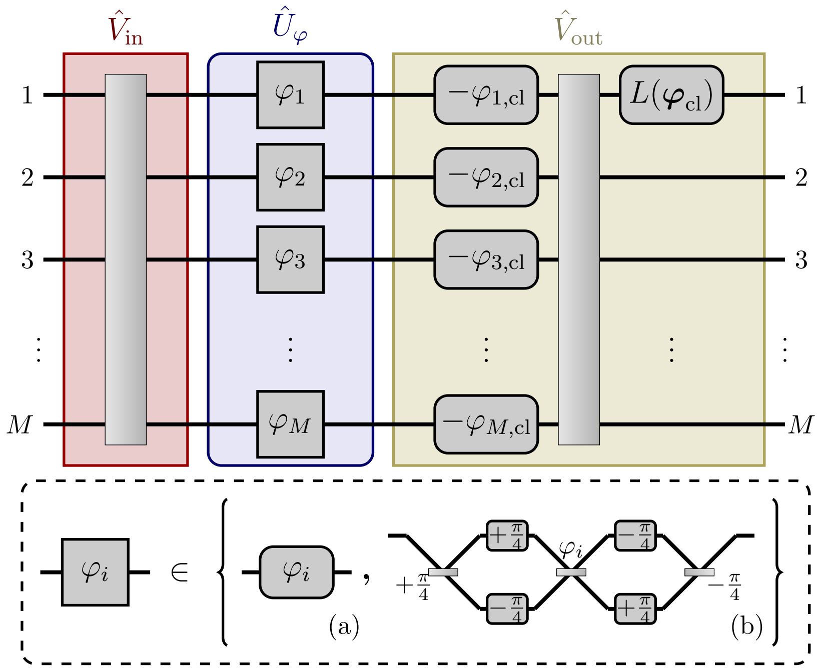

Figure 5. Network for the estimation at the Heisenberg limit of any linear combination of M parameters with positive weights, as shown in Equation (55) [37]. As shown in the lower panel, each parameter φ_i can either be (a) an optical phase acquired through a single-mode phase-shift, or (b) the reflectivity of a lossless beam-splitter. The network (b) is given in Equation (56), and its purpose is to transform the reflectivity into an optical phase. Two auxiliary stages $\hat{V}_{\text{in}}$ and $\hat{V}_{\text{out}}$ are employed, whose purpose is to distribute the probe according to the weights in the linear combination (see Equation (59)), and to refocus the probe into the first output port of the network (see Equation (60)).

With reference to Figure 5, we now describe the network $\hat{U}_{\boldsymbol{\varphi}}$ affected by M unknown parameters $\boldsymbol{\varphi}$. The parameters $\boldsymbol{\varphi}$ act in parallel, in the sense that the i-th parameter φ_i only affects the i-th mode of the network, and they can equivalently be the magnitudes of a phase-shift $U_{\text{PS}}(\varphi_i) = \mathrm{e}^{\mathrm{i}\varphi_i}$ or the reflectivities of beam-splitters $U_{\text{BS}}(\varphi_i) = \mathrm{e}^{\mathrm{i}\varphi_i \sigma_2}$, with σ_2 the second Pauli matrix. For each unknown beam-splitter, a 2-channel passive and linear network is employed in $\hat{U}_{\boldsymbol{\varphi}}$, whose purpose is to transform the reflectivity into a relative phase between the two ports of the beam-splitter (see the panel (b) in Figure 5). This can be done through the network

$$VU_{\mathrm{BS}}(\varphi)V^{\dagger} = U_{\mathrm{PS}}(\varphi) \equiv \begin{pmatrix} \mathrm{e}^{\mathrm{i}\varphi} & 0 \\ 0 & \mathrm{e}^{-\mathrm{i}\varphi} \end{pmatrix}, \tag{56}$$

with

$$V = U_{\mathrm{BS}}\left(\frac{\pi}{4}\right) U_{\mathrm{PS}}\left(\frac{\pi}{4}\right). \tag{57}$$

We also remark that, despite each mode of $\hat{U}_{\boldsymbol{\varphi}}$ with an unknown beam-splitter is practically composed of two separated channels, the overall network $VU_{\mathrm{BS}}(\varphi)V^{\dagger}$ in Equation (56) acts as a phase-shift in each of the two channels, namely the two modes are not mixed. Since this local network $VU_{\mathrm{BS}}(\varphi)V^{\dagger}$ is only fed through a single input port, it essentially acts as a single-channel phase shift of magnitude $\pm\varphi$ on the probe, depending on which of the two arms are employed. The overall network $\hat{U}_{\boldsymbol{\varphi}}$ can be thus described with the unitary matrix

$$U_{\boldsymbol{\varphi}} = \mathrm{diag}(\mathrm{e}^{\mathrm{i}\varphi_1}, \ldots, \mathrm{e}^{\mathrm{i}\varphi_M}), \tag{58}$$

regardless of the nature of the parameters $\boldsymbol{\varphi}$, whether they are phase magnitudes or beam-splitter reflectivities.

The input auxiliary stage is a M-channel generalized beam-splitter, which scatters the probe injected in the first input port into each of the M channels of $\hat{U}_{\boldsymbol{\varphi}}$ according to the weights $\boldsymbol{w}$. Specifically, the unitary matrix V_{in} representing the input network is chosen so that

$$|(V_{\mathrm{in}})_{i1}|^2 = w_i. \tag{59}$$

Noticeably, this constraint is φ-independent, i.e., $\hat{V}_{\mathrm{in}}$ does not need to be optimized after a prior coarse estimation of $\boldsymbol{\varphi}$. The output auxiliary network $\hat{V}_{\mathrm{out}}$ can be though as been composed of three separate stages. First, a phase shift of magnitude $-\varphi_{\mathrm{cl,i}}$ is applied to the i-th channel, where $\varphi_{\mathrm{cl,i}}$ is a coarse estimate of the prior classical measurement of φ_i, so that the error committed $\delta\varphi_i = \varphi_i - \varphi_{\mathrm{cl,i}}$ is of order $1/\sqrt{N}$. Then, a second generalized beam-splitter which does not depend on $\boldsymbol{\varphi}$ is in place, whose purpose is to invert the action of $\hat{V}_{\mathrm{in}}$ and thus refocusing the probe into the first channel of the network. Finally, a phase-shift of magnitude $L(\boldsymbol{\varphi}_{\mathrm{cl}})$ is applied before the homodyne detection at the first output port. The overall action of the network $\hat{V}_{\mathrm{out}}$ is thus described by the unitary matrix

$$V_{\mathrm{out}} = U_{\mathrm{PS}}(L(\boldsymbol{\varphi}_{\mathrm{cl}}), 0, \ldots, 0)\, V_{\mathrm{in}}^{\dagger}\, U_{\mathrm{PS}}(-\boldsymbol{\varphi}_{\mathrm{cl}}), \tag{60}$$

where we denoted with $U_{\mathrm{PS}}(\lambda_1, \ldots, \lambda_l) = \mathrm{diag}(\mathrm{e}^{\mathrm{i}\lambda_1}, \ldots, \mathrm{e}^{\mathrm{i}\lambda_l})$.

With this setup, the probability amplitude $(u_{\boldsymbol{\varphi}})_{11} = (V_{\mathrm{out}} U_{\boldsymbol{\varphi}} V_{\mathrm{in}})_{11}$ found in Equation (21) can be easily evaluated through Equations (58)–(60), and it reads

$$(u_{\boldsymbol{\varphi}})_{11} = \mathrm{e}^{\mathrm{i}L(\boldsymbol{\varphi}_{\mathrm{cl}})} \sum_{i=1}^{M} w_i \mathrm{e}^{\mathrm{i}\delta\varphi_i} = \mathrm{e}^{\mathrm{i}L(\boldsymbol{\varphi}_{\mathrm{cl}})} \left(1 + \sum_{i=1}^{M} \mathrm{i} w_i \delta\varphi_i - \frac{1}{2}\sum_{i=1}^{M} w_i \delta\varphi_i^2\right) + \mathcal{O}(\delta\boldsymbol{\varphi})^3, \tag{61}$$

where we can neglect the term $\mathcal{O}(\delta\boldsymbol{\varphi})^3$, since it is of order $N^{-3/2}$ and thus beyond the Heisenberg limit resolution we can achieve. We can now evaluate the transition probability $P_{\boldsymbol{\varphi}} = |u_{\boldsymbol{\varphi}}|^2_{11}$ through Equation (61)

$$P_{\boldsymbol{\varphi}} = \left|\mathrm{e}^{\mathrm{i}L(\boldsymbol{\varphi}_{\mathrm{cl}})} \sum_{i=1}^{M} w_i \mathrm{e}^{\mathrm{i}\delta\varphi_i}\right|^2 = 1 + \left(\sum_{i=1}^{M} w_i \delta\varphi_i\right)^2 - \sum_{i=1}^{M} w_i \delta\varphi_i^2 + \mathcal{O}(\delta\boldsymbol{\varphi})^3 \tag{62}$$

which clearly satisfies condition (29b), so that the Heisenberg limit can be achieved in the estimation of the complex phase $f(\boldsymbol{\varphi}) = \arg(u_{\boldsymbol{\varphi}})_{11}$. We can thus evaluate, comparing

Equations (29b) and (62), the factor ℓ which enters in the Fisher information shown in Equation (33),

$$\ell = \left(\sum_i w_i k_i\right)^2 - \sum_i w_i k_i^2, \tag{63}$$

where we supposed that $\delta\varphi_i = k_i/\sqrt{N}$ with k_i independent of N. From Equation (61) we obtain

$$f(\boldsymbol{\varphi}) = L(\boldsymbol{\varphi}_{\mathrm{cl}}) + \sum_{i=1}^{M} w_i \delta\varphi_i + \mathcal{O}(\delta\boldsymbol{\varphi})^3 = L(\boldsymbol{\varphi}) + \mathcal{O}(\delta\boldsymbol{\varphi})^3, \tag{64}$$

so that it is possible to recover the linear combination (55) at the Heisenberg limit from the estimation of $f(\boldsymbol{\varphi})$. We notice in fact that $L(\boldsymbol{\varphi})$ in (55) and the complex phase $f(\boldsymbol{\varphi})$ in (64) are equal up to a quantity of order $\mathcal{O}(N^{-3/2})$, which is beyond the Heisenberg resolution, since the errors $\delta\boldsymbol{\varphi}$ in the classical prior estimation are of order $1/\sqrt{N}$.

In conclusion, we will discuss about two different features of this protocol. First, we notice from Equation (56) that the network employed to transform the reflectivity φ of the beam-splitters into phase-shifts inverts the sign of the parameter in the second channel. This means that, if the portion of the probe which has been scattered into this network is injected in its second port, it acquires a phase $-\varphi$. It is then possible to exploit this behaviour to estimate a linear combination which admits negative weights for the reflectivities of the beam-splitters, employing the same protocol described in this section with the only precaution to invert the sign of the classical estimation φ_{cl} of this parameter. Second, it is possible to further generalize the local networks in each mode of $\hat{U}_{\boldsymbol{\varphi}}$. In fact, we can replace the networks in the panel of Figure 5 with a generic m-channel network $\hat{V}$ and, provided that the probe comes out of this local network from a single output channel—i.e., the network acts as a phase-shift on the probe— the same results found in this section still apply. Not only, in Appendix G we show that it is not necessary that all the photons come out from a single output port of each local network: instead, it suffices that a similar refocusing condition to the one in Equation (29b) is satisfied locally by each $\hat{V}$. It is thus possible to conceive, for example, a scheme that allows to reach the Heisenberg limit in the estimation of linear combination of parameters and functions of parameters, if we choose as $\hat{V}$ the whole network described in the example in Section 4.4.1.

5. Conclusions

The quantum metrological revolution is, at this moment in time, exciting and thrilling. The recent advances in the field of quantum mechanics have been incessantly stimulating the development of increasingly ingenious and innovative technologies which exploit the laws and rules of the microscopic world. Ranging from computing to biology, medicine, cosmology, imaging, sensing, cryptography and neural networks, quantum technologies appear to consistently outperform their classical counterparts. In particular, the fields of quantum sensing and quantum metrology propose schemes for the estimation of physical properties, such as lengths, time intervals, temperatures, and more, achieving enhanced levels of precision. In particular, the field of distributed Gaussian metrology, which studies strategies employing Gaussian states for the estimation of unknown parameters distributed arbitrarily on a passive and linear network, has recently witnessed some improvements in overcoming the challenge of adaptivity of the network, usually found in Gaussian and, more in general, quantum-enhanced schemes. In fact, as we briefly summarize in Section 2, employing a single squeezed vacuum state and performing homodyne detection at a single output port, it has been shown that it is possible to achieve the elusive Heisenberg-scaling sensitivity by only adding a single auxiliary network, whose purpose is to refocus the probe into the only observed channel. This auxiliary network, which introduced degrees of freedom in the interferometer, can be engineered with only a classical prior knowledge on the parameter.

The main focus of this review was then put on discussing the role and the effect of further degrees of freedom that can be added in the network, for example employing a second, non-optimized auxiliary stage. We have shown in Section 3 that this auxiliary stage leaves essentially unchanged the precision achieved by the setup, especially for networks with a large number of channel, and we have discussed how this result allows us to assume that the constant factor multiplying the Heisenberg scaling of the precision can be controlled and is typically far from zero. In Section 4 we demonstrated that it is possible to exploit this ability to manipulate how the information on multiple parameters encoded in a passive and linear network. We have seen that this results in the possibility to estimate functions of the unknown parameters which we can manipulate by acting on these further degrees of freedom through the auxiliary stages. The advantage of estimating directly functions of parameters lies in the fact that it allows us to save resources which would otherwise be employed to estimate singularly and at a high precision each parameter. In this way, both linear and non-linear functions can be estimated at the Heisenberg limit, and two example have been proposed.

Funding: This work was partially supported by the Office of Naval Research Global (N62909-18-1-2153). P.F. is partially supported by Istituto Nazionale di Fisica Nucleare (INFN) through the project QUANTUM, and by the Italian National Group of Mathematical Physics (GNFM-INdAM).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Acknowledgments: We thank Giovanni Gramegna and Frank A. Narducci for useful discussions.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A. Typicality of Gaussian Metrology

In this appendix, we will derive the statistical results discussed in Section 3 regarding the pre-factor $f(U, G_\varphi)$ appearing in the Fisher information in Equation (6). First, we will obtain the average of the pre-factor shown in Equation (18), employing results of computation of averages over the unitary group. Then, we will apply standard results on concentration of measure to derive the result in Equation (20).

Appendix A.1. Derivation of the Average of the Pre-Factor

Denoting with $\mathcal{P}$ the Haar probability measure defined on the unitary group $\mathrm{U}(M)$ of the $M \times M$ unitary matrices, it is possible to define the average of a given function $f : \mathrm{U}(M) \to \mathbb{C}$

$$\mathbb{E}[f(U)] \equiv \mathbb{E}_{\mathcal{P}}[f(U)] = \int f(U) \mathrm{d}\mathcal{P}(U). \tag{A1}$$

In order to derive the average $\mathbb{E}[f(U, G_\varphi)]$ in Equation (13)

$$f(U, G_\varphi) = (U^\dagger G_\varphi U)^2_{11}, \tag{A2}$$

we are interested only in the moments of the matrix elements U_{ij} up to the fourth orders, i.e., the averages of powers of the matrix elements and their complex conjugates. For random choices of the unitary matrix $U \in \mathrm{U}(M)$, the only non-vanishing moments up to the fourth order of the elements U_{ij} are given by [61]

$$\mathbb{E}[|U_{ij}|^2] = \frac{1}{M} \qquad (1 \le i, j \le M), \tag{A3a}$$

$$\mathbb{E}[|U_{ij}|^4] = \frac{2}{M(M+1)} \qquad (1 \le i, j \le M), \tag{A3b}$$

$$\mathbb{E}[|U_{ij}|^2 |U_{kj}|^2] = \frac{1}{M(M+1)} \qquad (i \neq k), \tag{A3c}$$

$$\mathbb{E}[|U_{ij}|^2|U_{il}|^2] = \frac{1}{M(M+1)} \qquad (j \neq l), \tag{A3d}$$

$$\mathbb{E}[|U_{ij}|^2|U_{kl}|^2] = \frac{1}{M^2-1} \qquad (i \neq k,\ j \neq l), \tag{A3e}$$

$$\mathbb{E}[U_{ij}U_{kl}U_{il}^*U_{kj}^*] = -\frac{1}{M(M^2-1)} \qquad (i \neq k,\ j \neq l). \tag{A3f}$$

The results shown in Equation (A3) can be conveniently expressed in two compact formulas [62] as:

$$\mathbb{E}[U_{ij}U_{kl}^*] = \frac{\delta_{ik}\delta_{jl}}{M} \tag{A4a}$$

$$\begin{aligned}\mathbb{E}[U_{ij}U_{kl}U_{mn}^*U_{pq}^*] = {} & \frac{\delta_{im}\delta_{jn}\delta_{kp}\delta_{lq} + \delta_{ip}\delta_{jq}\delta_{km}\delta_{ln}}{M^2-1} \\ & - \frac{\delta_{im}\delta_{jq}\delta_{kp}\delta_{ln} + \delta_{ip}\delta_{jn}\delta_{km}\delta_{lq}}{M(M^2-1)}\end{aligned} \tag{A4b}$$

Employing the formulas in Equations (A4), we are able to derive the averages in the main text in Equations (18)–(19). In fact, given a generic $M \times M$ complex matrix A, employing the result in Equation (A4a), we have

$$\begin{aligned}\mathbb{E}[(U^\dagger A U)_{ij}] &= \sum_{k,l} \mathbb{E}[U_{ik}^\dagger A_{kl} U_{lj}] \\ &= \sum_{k,l} A_{kl}\mathbb{E}[U_{ki}^* U_{lj}] \\ &= \sum_{k,l} A_{kl}\frac{\delta_{kl}\delta_{ij}}{M} \\ &= \frac{\mathrm{Tr}(A)}{M}\delta_{ij},\end{aligned} \tag{A5}$$

while employing the formula in Equation (A4), we have

$$\begin{aligned}\mathbb{E}[(U^\dagger A U)_{ij}^2] &= \sum_{k,l,m,n} \mathbb{E}[U_{ik}^\dagger A_{kl} U_{lj} U_{im}^\dagger A_{mn} U_{nj}] \\ &= \sum_{k,l,m,n} A_{kl} A_{mn} \mathbb{E}[U_{lj} U_{nj} U_{mi}^* U_{ki}^*] \\ &= [\mathrm{Tr}\left(A^2\right) + \mathrm{Tr}(A)^2]\left(\frac{1}{M^2-1} - \frac{1}{M(M^2-1)}\right)\delta_{ij} \\ &= \frac{\mathrm{Tr}(A^2) + \mathrm{Tr}(A)^2}{M(M+1)}\delta_{ij}\end{aligned} \tag{A6}$$

The expressions in Equations (A5) and (A6) reduce to the equalities in Equations (18) and (19) for $A = G_\varphi$ and $i = j = 1$, respectively.

Appendix A.2. Derivation of the Typicality Results

To show how to derive the result in Equation (20), we start from a standard result on concentration of measure in high-dimensional spaces known as Levy's Lemma

Theorem A1. *Let $f : \mathbb{S}^{n-1} \to \mathbb{R}$ be a function defined over the unit euclidean sphere*

$$\mathbb{S}^{n-1} = \left\{ \boldsymbol{x} \in \mathbb{R}^n \middle| \sum_{k=1}^{n} x_k^2 = 1 \right\} \tag{A7}$$

endowed with the uniform probability measure $\mathcal{P}$. Denote with L the Lipschitz constant of the function, such that

$$|f(\boldsymbol{x}) - f(\boldsymbol{y})| \leq L\|x - y\|_2, \tag{A8}$$

for all $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{S}^{n-1}$, where $\|\boldsymbol{x}\|_2 = \sqrt{\sum_{k=1}^{n} x_k^2}$ is the Euclidean norm. Then:

$$\mathcal{P}(|f - \mathbb{E}[f]| \geq \varepsilon) \leq 2\mathrm{e}^{-\frac{n\varepsilon^2}{CL^2}}, \tag{A9}$$

where C is some positive constant which can be taken to be $C = 9\pi^3$ [63,64].

To prove the concentration result in Equation (20), we want to apply Theorem A1 to our case. Thus, we first need to compute the Lipschitz constant L associated with the pre-factor in Equation (13). To do so, we first notice that $f(U, G_\varphi)$ can be thought as a function defined on the real unit sphere $\mathbb{S}^{n-1}$. In fact, it can be written as

$$f(U, G_\varphi) = \left(\sum_{j=1}^{M} |u_j|^2 g_j \right)^2. \tag{A10}$$

where $\boldsymbol{u}$ is the complex M-dimensional vector given by $\boldsymbol{u} = V_\varphi^\dagger U \boldsymbol{e}$ with $\boldsymbol{e} = (1, 0, \ldots, 0)^T \in \mathbb{C}^M$, as shown in Equation (15), with $V_\varphi^\dagger$ being the matrix whose columns are the eigenvectors of G_φ. Since only the squared moduli $|u_j|^2$ of this complex vector appear in Equation (A10), we can consider the pre-factor $f(U, G_\varphi)$ as a function of a real vector $\boldsymbol{x} \in \mathbb{R}^{2M}$ whose $2M$ components are defined by

$$x_{2j-1} = \operatorname{Re} u_j, \qquad x_{2j} = \operatorname{Im} u_j, \qquad j = 1, \ldots, M. \tag{A11}$$

The unitarity constraint $\sum_{j=1}^{M} |u_j|^2 = 1$ translates into $\sum_{j=1}^{2M} x_j^2 = 1$, so that $\boldsymbol{x} \in \mathbb{S}^{2M-1}$, the unit sphere sitting inside $\mathbb{R}^{2M}$. We see then that the random factor in Equation (A10) can be thought as a function defined over the unit sphere $\mathbb{S}^{2M-1}$:

$$\begin{aligned} f(U, G_\varphi) &= \left(\sum_{j=1}^{M} |u_j|^2 g_j \right)^2 \\ &= \left(\sum_{j=1}^{M} (x_{2j-1}^2 + x_{2j}^2) g_j \right)^2 \qquad \text{(A12)} \\ &= \left(\boldsymbol{x}^T \widetilde{G} \boldsymbol{x} \right)^2 =: f(\boldsymbol{x}), \qquad \text{(A13)} \end{aligned}$$

where we have defined the diagonal matrix $\widetilde{G} = \operatorname{diag}(\widetilde{g})$ with $\widetilde{g} = (g_1, g_1, \ldots, g_M, g_M) \in \mathbb{R}^{2M}$. We can now estimate the Lipschitz constant L of the function $f(\boldsymbol{x})$ to apply Theorem A1. To this aim, we evaluate the gradient of f, which is given by:

$$\boldsymbol{\nabla} f(x) = 4(\boldsymbol{x}^T \widetilde{G} \boldsymbol{x}) \widetilde{G} \boldsymbol{x}. \tag{A14}$$

The Lipschitz constant for f can be then obtained as

$$L = \max_{\boldsymbol{x} \in \mathbb{S}^{2M-1}} \|\boldsymbol{\nabla} f(\boldsymbol{x})\|_2 = 4\|G_\varphi\|^2, \tag{A15}$$

since

$$
\begin{aligned}
\|\nabla f(\boldsymbol{x})\|_2 &= \sqrt{[\nabla f(\boldsymbol{x})]^T[\nabla f(\boldsymbol{x})]} \\
&= 4|\boldsymbol{x}^T\widetilde{G}\boldsymbol{x}|\sqrt{\boldsymbol{x}^T\widetilde{G}^2\boldsymbol{x}} \\
&\leq 4\|\widetilde{G}\|^2 \\
&= 4\left\|G_\varphi\right\|^2
\end{aligned} \tag{A16}
$$

where we used the fact that $|\boldsymbol{x}^T\widetilde{G}\boldsymbol{x}| \leq \|\widetilde{G}\|$ and $\boldsymbol{x}^T\widetilde{G}^2\boldsymbol{x} = \|\widetilde{G}\boldsymbol{x}\|^2 \leq \|\widetilde{G}\|^2$, while in the last equality we used the fact that $\|\widetilde{G}\| = \left\|G_\varphi\right\| = \max_i g_i$. The value $\|\nabla f(\boldsymbol{x})\|_2 = 4\left\|G_\varphi\right\|^2$ can be obtained with $\boldsymbol{x} = \boldsymbol{e}_i$, supposing that g_i is the eigenvalue with highest absolute value. The value of $L = 4\left\|G_\varphi\right\|^2$ in Equation (A15) can then be used when applying Theorem A1, with $n = 2M$, to finally prove the result in Equation (20).

Appendix B. Probability Distributions from Homodyne Measurements

In this appendix, we will obtain the probability density functions which governs the outcomes of homodyne detections performed on a single-mode squeezed state $|\boldsymbol{r}\rangle = \hat{S}(\boldsymbol{r})|\mathrm{vac}\rangle$ injected in the first port of a M-channel passive and linear network $\hat{U}$, with $\boldsymbol{r} = (r, 0, \ldots, 0)$, so that $N = \sinh^2 r$, with $r > 0$, and $N = N_{\mathrm{S}}$ average number of photons. The state $|\boldsymbol{r}\rangle$ is a Gaussian state which is completely described, through its Wigner distribution [29–32], by its covariance matrix $\Gamma = \mathrm{diag}(\mathrm{e}^{2r}, 1\ldots, 1, \mathrm{e}^{-2r}, 1\ldots, 1)/2$. Moreover, passive and linear networks preserve the Gaussian nature of $|\boldsymbol{r}\rangle$ [29–32]. For this reason, we will first obtain in generality the final covariance matrix Γ_U of the state $\hat{U}|\boldsymbol{r}\rangle$, which define its Wigner distribution. Then, we will marginalize the Wigner distribution associated with the first channel—since it is the only observed port—rotated by the symplectic and orthogonal matrix

$$
R(\theta) = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}, \tag{A17}
$$

where θ is the local oscillator phase of the homodyne. In doing so, we will be able to obtain the expression of the variance $\sigma_{\boldsymbol{\varphi}}$ in Equation (23).

The covariance matrix Γ_U of the state $\hat{U}|\boldsymbol{r}\rangle$ is given by the transformation $\Gamma_U = R\Gamma R^{\mathrm{T}}$, with

$$
R = \begin{pmatrix} \mathrm{Re}(U) & -\mathrm{Im}(U) \\ \mathrm{Im}(U) & \mathrm{Re}(U) \end{pmatrix}, \tag{A18}
$$

orthogonal and symplectic matrix representing the rotation in the phase space generated by the network $\hat{U}$ [29–32]. A straightforward calculation shows that

$$
\Gamma_U = R\Gamma R^{\mathrm{T}} = \begin{pmatrix} \Delta X^2 & \Delta XP \\ \Delta XP^T & \Delta P^2 \end{pmatrix}. \tag{A19}
$$

where we have defined the $M \times M$ matrices

$$
\begin{aligned}
\Delta X^2 &\equiv \frac{1}{2}\Big[\mathrm{Re}[U]\mathrm{e}^{2\mathcal{R}}\,\mathrm{Re}[U^\dagger] - \mathrm{Im}[U]\mathrm{e}^{-2\mathcal{R}}\,\mathrm{Im}[U^\dagger]\Big] \\
&= \frac{1}{2}\Big[\mathrm{Re}[U\cosh(2\mathcal{R})U^\dagger] + \mathrm{Re}[U\sinh(2\mathcal{R})U^{\mathrm{T}}]\Big],
\end{aligned} \tag{A20a}
$$

$$
\begin{aligned}
\Delta P^2 &\equiv \frac{1}{2}\Big[-\mathrm{Im}[U]\mathrm{e}^{2\mathcal{R}}\,\mathrm{Im}[U^\dagger] + \mathrm{Re}[U]\mathrm{e}^{-2\mathcal{R}}\,\mathrm{Re}[U^\dagger]\Big] \\
&= \frac{1}{2}\Big[\mathrm{Re}[U\cosh(2\mathcal{R})U^\dagger] - \mathrm{Re}[U\sinh(2\mathcal{R})U^{\mathrm{T}}]\Big],
\end{aligned} \tag{A20b}
$$

$$
\begin{aligned}
\Delta XP &\equiv \frac{1}{2}\Big[-\mathrm{Re}[U]\mathrm{e}^{2\mathcal{R}}\,\mathrm{Im}[U^\dagger] - \mathrm{Im}[U]\mathrm{e}^{-2\mathcal{R}}\,\mathrm{Re}[U^\dagger]\Big] \\
&= \frac{1}{2}\Big[-\mathrm{Im}[U\cosh(2\mathcal{R})U^\dagger] + \mathrm{Im}[U\sinh(2\mathcal{R})U^{\mathrm{T}}]\Big].
\end{aligned} \tag{A20c}
$$

The 2×2 reduced covariance matrix Γ_U^1 of the first mode reads

$$\Gamma_U^1 = \begin{pmatrix} (\Delta X^2)_{11} & (\Delta XP)_{11} \\ (\Delta XP)_{11} & (\Delta P^2)_{11} \end{pmatrix}. \tag{A21}$$

Our final step is to recover the variance of the quadrature $\hat{x}_\theta$. In order to do that, we employ the orthogonal and symplectic matrix $R(\theta)$ in Equation (A17), representing the action of a phase-shift $e^{-i\theta}$, namely a clock-wise rotation of an angle θ in the first mode phase-space. The variance $\sigma_{\boldsymbol{\varphi}}^2$ in (23) is finally obtained by a direct computation, by recalling the parametrization $(U_{\boldsymbol{\varphi}})_{11} = \sqrt{P_{\boldsymbol{\varphi}}} e^{if(\boldsymbol{\varphi})}$ in Equation (21)

$$\begin{aligned} \sigma_{\boldsymbol{\varphi}}^2 &= (O_\theta \Gamma_U O_\theta^T)_{11} \\ &= \frac{1}{2} + P_{\boldsymbol{\varphi}}(\sinh^2(r) + \cos(2f(\boldsymbol{\varphi}) - 2\theta)\cosh(r)\sinh(r)). \end{aligned} \tag{A22}$$

recovering the expression in Equation (23).

Appendix C. Asymptotic Analysis of Gaussian Metrology

We here evaluate the asymptotic expressions of the Fisher information matrix in Equation (30), showing that the conditions (29) yield the Heisenberg-scaling sensitivity for the estimation of $f(\boldsymbol{\varphi})$. As shown in Equations (26) and (27), the dependence of the variance

$$\sigma_{\boldsymbol{\varphi}}^2 = \frac{1 - P_{\boldsymbol{\varphi}}}{2} + \frac{P_{\boldsymbol{\varphi}}}{2}\left[\cosh(2r) + \cos(2\gamma_{\boldsymbol{\varphi}} - 2\theta)\sinh(2r)\right], \tag{A23}$$

on the parameters $\boldsymbol{\varphi}$ only appears through the transition probability $P_{\boldsymbol{\varphi}}$ and the acquired complex phase $\gamma_{\boldsymbol{\varphi}}$

$$\nabla\sigma_{\boldsymbol{\varphi}}^2 = (\partial_P \sigma_{\boldsymbol{\varphi}}^2)\nabla P_{\boldsymbol{\varphi}} + (\partial_f \sigma_{\boldsymbol{\varphi}}^2)\nabla f(\boldsymbol{\varphi}), \tag{A24}$$

where ∂_P and ∂_f represent the differentiation with respect of $P_{\boldsymbol{\varphi}}$ and $f(\boldsymbol{\varphi})$, and

$$\partial_P \sigma_{\boldsymbol{\varphi}}^2 = \frac{1}{2}(-1 + \cosh(2r) + \cos(2f(\boldsymbol{\varphi}) - 2\theta)\sinh(2r)) \tag{A25a}$$

$$\partial_f \sigma_{\boldsymbol{\varphi}}^2 = -P_{\boldsymbol{\varphi}}\sin(2f(\boldsymbol{\varphi}) - 2\theta)\sinh(2r). \tag{A25b}$$

As discussed in Section 4.2, to achieve the HL, some conditions must be imposed so that the variance in Equation (A23) does not grow with N. The only option to do that without ruining the sensitivity of the setup is requesting that $\gamma_{\boldsymbol{\varphi}} - \theta \simeq \pi/2$, as we can see from Equation (A23). In particular, we impose condition (29a), and evaluate the variance in (A23) and its gradient (A24) in the large N limit

$$\begin{aligned} \sigma_{\boldsymbol{\varphi}}^2 &= \frac{1}{2} + NP_{\boldsymbol{\varphi}}\left(1 - \cos\left(\frac{2k}{N}\right)\sqrt{1 + \frac{1}{N}}\right) \\ &= \frac{1}{2} + NP_{\boldsymbol{\varphi}}\left(1 - \left(1 - \frac{2k^2}{N^2}\right)\left(1 + \frac{1}{2N} - \frac{1}{8N^2}\right)\right) + \mathcal{O}\left(\frac{1}{N^2}\right) \\ &= \frac{1 - P_{\boldsymbol{\varphi}}}{2} + P_{\boldsymbol{\varphi}}\left(\frac{2k^2}{N} + \frac{1}{8N}\right) + \mathcal{O}\left(\frac{1}{N^2}\right), \end{aligned} \tag{A26}$$

$$\begin{aligned} \nabla\sigma_{\boldsymbol{\varphi}}^2 &= N\nabla P_{\boldsymbol{\varphi}}\left(1 - \cos\left(\frac{2k}{N}\right)\sqrt{1 + \frac{1}{N}}\right) + 2NP_{\boldsymbol{\varphi}}\nabla f(\boldsymbol{\varphi})\sin\left(\frac{2k}{N}\right)\sqrt{1 + \frac{1}{N}} \\ &= N\nabla P_{\boldsymbol{\varphi}}\left(1 - \left(1 + \frac{1}{2N}\right)\right) + 2NP_{\boldsymbol{\varphi}}\nabla f(\boldsymbol{\varphi})\frac{2k}{N} + \mathcal{O}\left(\frac{1}{N}\right) \\ &= -\frac{1}{2}\nabla P_{\boldsymbol{\varphi}} + 4kP_{\boldsymbol{\varphi}}\nabla f(\boldsymbol{\varphi}) + \mathcal{O}\left(\frac{1}{N}\right). \end{aligned} \tag{A27}$$

Then, by imposing condition (29b) on $P_{\boldsymbol{\varphi}}$, we get

$$\sigma_{\boldsymbol{\varphi}}^2 = \left(2k^2 + \frac{1}{8} + \frac{\ell}{2}\right)\frac{1}{N} + \mathcal{O}\left(\frac{1}{N^2}\right), \tag{A28a}$$

$$\nabla\sigma_{\boldsymbol{\varphi}}^2 = 4k\nabla f(\boldsymbol{\varphi}) + \mathcal{O}\left(\frac{1}{N}\right). \tag{A28b}$$

By substituting the expressions in Equations (A28) into the Fisher information matrix in Equation (25), we can finally evaluate its asymptotic expression

$$\mathcal{F}(\boldsymbol{\varphi}) \sim 8\varrho(k,\ell)N^2\big(\nabla f(\boldsymbol{\varphi})\big)\big(\nabla f(\boldsymbol{\varphi})\big)^{\mathrm{T}} \tag{A29}$$

with

$$\varrho(k,\ell) = \left(\frac{8k}{16k^2 + 1 + 4\ell}\right)^2 \tag{A30}$$

a positive and N-independent pre-factor.

Appendix D. Maximum-Likelihood Estimators for Gaussian Distributions

In this Appendix we will find the solution which maximizes the Likelihood function in Equations (34) and (35).

Due to the monotonicity of the logarithmic function, we can maximize the log-likelihood function, and thus obtain

$$\begin{aligned} 0 &= \frac{\partial}{\partial f}\ln\mathcal{L}(\boldsymbol{\varphi}|\vec{x})\Big|_{f=\tilde{f}_{\mathrm{MLE}}} \\ &= \frac{\partial}{\partial f}\sum_{j=1}^{\nu}\ln p(x_j|\boldsymbol{\varphi})\Big|_{f=\tilde{f}_{\mathrm{MLE}}} \\ &= \frac{\partial}{\partial f}\sum_{j=1}^{\nu}\left(-\frac{1}{2}\ln\sigma_{\boldsymbol{\varphi}}^2 - \frac{x_j^2}{2\sigma_{\boldsymbol{\varphi}}^2}\right)\Bigg|_{f=\tilde{f}_{\mathrm{MLE}}} \\ &= \left(\frac{\partial\sigma_{\boldsymbol{\varphi}}^2}{\partial f}\sum_{j=1}^{\nu}\left(-\frac{1}{2\sigma_{\boldsymbol{\varphi}}^2} + \frac{x_j^2}{2\sigma_{\boldsymbol{\varphi}}^4}\right)\right)\Bigg|_{f=\tilde{f}_{\mathrm{MLE}}}. \end{aligned} \tag{A31}$$

By assuming that $\partial\sigma_{\boldsymbol{\varphi}}^2/\partial f \neq 0$ the solution is given by the value of $f(\boldsymbol{\varphi})$ that solves

$$\sigma_{\boldsymbol{\varphi}}^2 = \sigma^2(\vec{x}) \equiv \frac{1}{\nu}\sum_{j=1}^{\nu}x_j^2, \tag{A32}$$

where $\sigma_{\boldsymbol{\varphi}}^2$ is given by (23), with $P_{\boldsymbol{\varphi}}$ given by (29b) and θ being phase of the local oscillator. Thus the estimator is ultimately given by

$$\tilde{f}_{\mathrm{MLE}}(\boldsymbol{x}) = \theta + \frac{1}{2}\left(2n\pi \pm \arccos\left(\frac{(2S(\boldsymbol{x})^2 - 1) - 2P_{\boldsymbol{\varphi}}\sinh^2 r}{2P_{\boldsymbol{\varphi}}\sinh r\cosh r}\right)\right), \tag{A33}$$

as shown in Equation (37).

Appendix E. Derivation of the Cramér-Rao Bound for Singular Fisher Information Matrix

In this appendix we will show that the Fisher information matrix in Equation (30)

$$\mathcal{F}(\boldsymbol{\varphi}) = 8\varrho(k,\ell)N^2\big(\nabla f(\boldsymbol{\varphi})\big)\big(\nabla f(\boldsymbol{\varphi})\big)^{\mathrm{T}}, \tag{A34}$$

which, as discussed in Section 4.2, is a matrix of rank$(\mathcal{F}(\boldsymbol{\varphi})) = 1$, whose only non-vanishing eigenvalue is given by

$$\lambda = 8\varrho(k,\ell)N^2|\nabla f(\boldsymbol{\varphi})|^2, \tag{A35}$$

yields the Cramér-Rao bound found in Equation (33) for the estimation of the function $f(\boldsymbol{\varphi})$.

The Cramér-Rao bound associated with the estimation of $f(\boldsymbol{\varphi})$ for non-invertible matrices can be written in terms of the Moore-Penrose pseudo-inverse $\mathcal{F}(\boldsymbol{\varphi})^+$ through the inequality [59]

$$\mathrm{Var}[\tilde{f}] \geqslant \frac{1}{\nu} H\mathcal{F}(\boldsymbol{\varphi})^+ H^{\mathrm{T}}, \tag{A36}$$

where

$$H = \big(\nabla f(\boldsymbol{\varphi})\big)^{\mathrm{T}} \equiv |\nabla f(\boldsymbol{\varphi})|\boldsymbol{v}^{\mathrm{T}}, \tag{A37}$$

is the gradient of the function $f(\boldsymbol{\varphi})$ which can be estimated with finite variance, which coincides with the eigenvector of $\mathcal{F}(\boldsymbol{\varphi})$ associated with the eigenvalue λ in Equation (A35), while $\boldsymbol{v}$ is the eigenvector normalized to the unit length. Since $\boldsymbol{v}$ is the only (normalized) eigenvector associated with λ, the pseudo-inverse $\mathcal{F}(\boldsymbol{\varphi})^+$ can be written as

$$\mathcal{F}(\boldsymbol{\varphi})^+ = \frac{1}{\lambda}\boldsymbol{v}\,\boldsymbol{v}^{\mathrm{T}} = \frac{1}{|\nabla f(\boldsymbol{\varphi})|^4}\frac{1}{8\varrho(k,\ell)N^2}\big(\nabla f(\boldsymbol{\varphi})\big)\big(\nabla f(\boldsymbol{\varphi})\big)^{\mathrm{T}}. \tag{A38}$$

We can thus evaluate the Cramér-Rao bound in Equation (A36)

$$\mathrm{Var}_{\boldsymbol{\varphi}}[\tilde{f}] \geq \frac{1}{\nu}|\nabla f(\boldsymbol{\varphi})|\boldsymbol{v}^{\mathrm{T}}\left(\frac{1}{\lambda}\boldsymbol{v}\,\boldsymbol{v}^{\mathrm{T}}\right)|\nabla f(\boldsymbol{\varphi})|\boldsymbol{v} = \frac{1}{\nu}\frac{1}{8\varrho(k,\ell)N^2} \tag{A39}$$

as displayed in Equation (33).

Appendix F. Analysis of the Transition Amplitude

We will here show that the transition amplitude

$$\begin{aligned}(u_{\boldsymbol{\varphi}})_{11} = \mathrm{e}^{\mathrm{i}\left(\alpha_1+\alpha_3+\frac{\delta\varphi_2+\delta\varphi_3}{2}\right)}\Bigg(&\cos\left(\frac{\delta\varphi_2-\delta\varphi_3}{2}\right)\sin 2\omega\cos\varphi_1\\ &+\cos\left(\Delta\alpha-\frac{\delta\varphi_2-\delta\varphi_3}{2}\right)\cos 2\omega\sin\varphi_1 + \mathrm{i}\sin\left(\Delta\alpha-\frac{\delta\varphi_2-\delta\varphi_3}{2}\right)\sin\varphi_1\Bigg),\end{aligned} \tag{A40}$$

shown in Equation (45) satisfies condition (29b) for the choices of ω satisfying Equation (43)

$$\omega = \frac{1}{2}\arctan\left(\frac{\cos(\varphi_{\mathrm{cl},1})}{\sin(\varphi_{\mathrm{cl},1})\cos\Delta\alpha}\right), \tag{A41}$$

and that the complex phase of $(u_{\boldsymbol{\varphi}})_{11}$ is the one given in Equations (46) and (47).

First, we notice that if condition (A41) holds, we can write

$$\sin 2\omega = \pm\frac{\cos(\varphi_{\mathrm{cl},1})}{\sqrt{\cos^2(\varphi_{\mathrm{cl},1}) + \sin^2(\varphi_{\mathrm{cl},1})\cos^2\Delta\alpha}} \tag{A42a}$$

$$\cos 2\omega = \pm\frac{\sin(\varphi_{\mathrm{cl},1})\cos\Delta\alpha}{\sqrt{\cos^2(\varphi_{\mathrm{cl},1}) + \sin^2(\varphi_{\mathrm{cl},1})\cos^2\Delta\alpha}}, \tag{A42b}$$

where the signs of both the right-hand expressions must be the same. We will only perform the calculation with both signs being positive, but the same steps can be done for the other case. In order to show that the probability of transition $P_{\boldsymbol{\varphi}} = \left|(u_{\boldsymbol{\varphi}})_{11}\right|^2$ satisfies condition (29b), we will evaluate $P_{\boldsymbol{\varphi}}$ in the case of perfect prior knowledge on the parameters—i.e., $\delta\varphi_i = \varphi_i - \varphi_{\mathrm{cl,i}} = 0$— and show that in this case $P_{\boldsymbol{\varphi}} = 1$. This is enough to show that condition (29b) is satisfied when the prior knowledge on the parameter is classical – namely $\delta\varphi_i = \mathcal{O}(N^{-1/2})$ – as discussed in Sec 4.3, since the first non-vanishing order in $\delta\boldsymbol{\varphi}_{\text{cl}}$ of $P_{\boldsymbol{\varphi}} - 1$ is $\mathcal{O}(\delta\boldsymbol{\varphi}_{\text{cl}})^2$. We thus employ the expressions in Equations (A42) with $\varphi_{\text{cl},1} = \varphi_1$ to evaluate

$$\begin{aligned}\left|(u_{\boldsymbol{\varphi}})_{11}\right|^2 &= (\sin 2\omega \cos\varphi_1 + \cos(\Delta\alpha)\cos 2\omega \sin\varphi_1)^2 + \sin^2(\Delta\alpha)\sin^2\varphi_1 \\ &= \frac{(\cos^2\varphi_1 + \sin^2\varphi_1\cos^2(\Delta\alpha))^{\cancel{2}}}{\cancel{\cos^2\varphi_1 + \sin^2\varphi_1\cos^2\Delta\alpha}} + \sin^2(\Delta\alpha)\sin^2\varphi_1 = 1. \end{aligned} \tag{A43}$$

To evaluate the complex phase $\arg(u_{\boldsymbol{\varphi}})_{11}$ in Equation (46), and in particular to prove the expression for Φ in Equation (47), we simply need to replace the condition on Equations (A42) into

$$\Phi = \arctan\left(\frac{\sin\left(\Delta\alpha - \frac{\delta\varphi_2 - \delta\varphi_3}{2}\right)\sin\varphi_1}{\cos\left(\frac{\delta\varphi_2-\delta\varphi_3}{2}\right)\sin 2\omega\cos\varphi_1 + \cos\left(\Delta\alpha - \frac{\delta\varphi_2-\delta\varphi_3}{2}\right)\cos 2\omega \sin\varphi_1}\right), \tag{A44}$$

from which we easily verify Equation (47).

Appendix G. Generalized Setup

In this appendix we will show that it is possible to employ more generic multi-channel local networks $\hat{U}_{i,\varphi_i}$ within the overall network $\hat{U}_{\varphi}$ in Figure 5, and still reaching the HL in the estimation of the linear combination

$$L(\boldsymbol{\varphi}) = \boldsymbol{w}\cdot\boldsymbol{\varphi} \tag{A45}$$

in Equation (55), if conditions (29) are satisfied. In particular, the parameters $\boldsymbol{\varphi}$ appearing in Equation (A45) will be in this case the phases acquired by the portion of the probe injected in each local network $\hat{U}_{i,\varphi_i}$. These phases can be in turn parameters distributed in each local network, or functions of parameters, as in the setup shown in Figures 1 and 4. We will show that the requirement that these local networks $\hat{U}_{i,\varphi_i}$ must met, in order to satisfy condition (29b), is

$$P_i \sim 1 - \frac{\ell_i}{N}, \qquad \ell_i \geq 0,\ i = 1,\ldots,m_2, \tag{A46}$$

similar to the global condition (29b), where P_i is the transition probability, associated with each local setup $\hat{U}_{i,\varphi_i}$, that a photon injected in the first channel of the i-th local network comes out from its upper channel, – i.e., $P_i = \left|U_{i,\varphi_i}\right|^2_{11}$. If conditions (A46) are satisfied, we can generalize the global transition amplitude in Equation (61) to

$$\begin{aligned}\chi_{\boldsymbol{\varphi}} &= \mathrm{e}^{\mathrm{i}L(\boldsymbol{\varphi}_{\text{cl}})}\sum_{i=1}^{M} w_i\sqrt{1-\frac{\ell_i}{N}}\mathrm{e}^{\mathrm{i}\delta\varphi_i} = \\ &= \mathrm{e}^{\mathrm{i}L(\boldsymbol{\varphi}_{\text{cl}})}\left(1 + \sum_{i=1}^{M}\mathrm{i}w_i\delta\varphi_i - \frac{1}{2}\sum_{i=1}^{M}w_i\left(\delta\varphi_i^2 + \frac{\ell_i}{N}\right)\right) + \mathcal{O}(N^{-3/2}),\end{aligned} \tag{A47}$$

where we made use of condition (A46) to write the transition amplitudes associated with each local network of $\hat{U}_{i,\varphi_i}$. Exploiting once again the requirement that prior estimations are classical, i.e., that $\delta\varphi_i = \mathcal{O}(N^{-1/2})$, we notice that the probability,

$$\begin{aligned}P_{\boldsymbol{\varphi}} &= \left|\mathrm{e}^{\mathrm{i}L(\boldsymbol{\varphi}cl)}\sum_{i=1}^{M}w_i\sqrt{1-\frac{\ell_i}{N}}\mathrm{e}^{\mathrm{i}\delta\varphi_i}\right|^2 = \\ &= 1 + \left(\sum_{i=1}^{M}w_i\delta\varphi_i\right)^2 - \sum_{i=1}^{M}w_i\left(\delta\varphi_i^2 + \frac{\ell_i}{N}\right) + \mathcal{O}(N^{-3/2})\end{aligned}$$

$$\equiv 1 - \frac{\ell}{N} + \mathcal{O}(N^{-3/2}) \tag{A48}$$

still satisfies condition (29b). Thus, the HL in the estimation of the total acquired phase shown in (A45) is still achieved for generic local networks satisfying the conditions (A46).

References

1. McConnell, R.; Low, G.H.; Yoder, T.J.; Bruzewicz, C.D.; Chuang, I.L.; Chiaverini, J.; Sage, J.M. Heisenberg scaling of imaging resolution by coherent enhancement. *Phys. Rev. A* **2017**, *96*, 051801. [CrossRef]
2. Unternährer, M.; Bessire, B.; Gasparini, L.; Perenzoni, M.; Stefanov, A. Super-resolution quantum imaging at the Heisenberg limit. *Optica* **2018**, *5*, 1150–1154. [CrossRef]
3. De Pasquale, A.; Stace, T.M. Quantum Thermometry. In *Thermodynamics in the Quantum Regime*; Springer: Cham, Switzerland, 2018; Volume 195, pp. 503–527. [CrossRef]
4. Seah, S.; Nimmrichter, S.; Grimmer, D.; Santos, J.P.; Scarani, V.; Landi, G.T. Collisional Quantum Thermometry. *Phys. Rev. Lett.* **2019**, *123*, 180602. [CrossRef] [PubMed]
5. Razzoli, L.; Ghirardi, L.; Siloi, I.; Bordone, P.; Paris, M.G.A. Lattice quantum magnetometry. *Phys. Rev. A* **2019**, *99*, 062330. [CrossRef]
6. Bhattacharjee, S.; Bhattacharya, U.; Niedenzu, W.; Mukherjee, V.; Dutta, A. Quantum magnetometry using two-stroke thermal machines. *New J. Phys.* **2020**, *22*, 013024. [CrossRef]
7. Aasi, J.; Abadie, J.; Abbott, B.P.; Abbott, R.; Abbott, T.D.; Abernathy, M.R.; Adams, C.; Adams, T.; Addesso, P.; Adhikari, R.X.; et al. Enhanced sensitivity of the LIGO gravitational wave detector by using squeezed states of light. *Nat. Photonics* **2013**, *7*, 613–619. [CrossRef]
8. Caves, C.M. Quantum-mechanical noise in an interferometer. *Phys. Rev. D* **1981**, *23*, 1693–1708. [CrossRef]
9. Bondurant, R.S.; Shapiro, J.H. Squeezed states in phase-sensing interferometers. *Phys. Rev. D* **1984**, *30*, 2548–2556. [CrossRef]
10. Wineland, D.J.; Bollinger, J.J.; Itano, W.M.; Moore, F.L.; Heinzen, D.J. Spin squeezing and reduced quantum noise in spectroscopy. *Phys. Rev. A* **1992**, *46*, R6797–R6800. [CrossRef]
11. Giovannetti, V.; Lloyd, S.; Maccone, L. Quantum-Enhanced Measurements: Beating the Standard Quantum Limit. *Science* **2004**, *306*, 1330–1336. [CrossRef]
12. Giovannetti, V.; Lloyd, S.; Maccone, L. Quantum Metrology. *Phys. Rev. Lett.* **2006**, *96*, 010401. [CrossRef] [PubMed]
13. Dowling, J.P. Quantum optical metrology—The lowdown on high-N00N states. *Contemp. Phys.* **2008**, *49*, 125–143. [CrossRef]
14. Paris, M.G. Quantum estimation for quantum technology. *Int. J. Quantum Inf.* **2009**, *7*, 125–137. [CrossRef]
15. Giovannetti, V.; Lloyd, S.; Maccone, L. Advances in quantum metrology. *Nat. Photonics* **2011**, *5*, 222–229. [CrossRef]
16. Lang, M.D.; Caves, C.M. Optimal Quantum-Enhanced Interferometry Using a Laser Power Source. *Phys. Rev. Lett.* **2013**, *111*, 173601. [CrossRef]
17. Tóth, G.; Apellaniz, I. Quantum metrology from a quantum information science perspective. *J. Phys. A Math. Theor.* **2014**. *47*, 424006. [CrossRef]
18. Erol, V.; Ozaydin, F.; Altintas, A.A. Analysis of Entanglement Measures and LOCC Maximized Quantum Fisher Information of General Two Qubit Systems. *Sci. Rep.* **2014**, *4*, 5422. [CrossRef]
19. Dowling, J.P.; Seshadreesan, K.P. Quantum Optical Technologies for Metrology, Sensing, and Imaging. *J. Light. Technol.* **2015**, *33*, 2359–2370. [CrossRef]
20. Czekaj, L.; Przysiężna, A.; Horodecki, M.; Horodecki, P. Quantum metrology: Heisenberg limit with bound entanglement. *Phys. Rev. A* **2015**, *92*, 062303. [CrossRef]
21. Ozaydin, F.; Altintas, A.A. Quantum Metrology: Surpassing the shot-noise limit with Dzyaloshinskii-Moriya interaction. *Sci. Rep.* **2015**, *5*, 16360. [CrossRef]
22. Szczykulska, M.; Baumgratz, T.; Datta, A. Multi-parameter quantum metrology. *Adv. Phys. X* **2016**, *1*, 621–639. [CrossRef]
23. Schnabel, R. Squeezed states of light and their applications in laser interferometers. *Phys. Rep.* **2017**, *684*, 1–51. [CrossRef]
24. Braun, D.; Adesso, G.; Benatti, F.; Floreanini, R.; Marzolino, U.; Mitchell, M.W.; Pirandola, S. Quantum-enhanced measurements without entanglement. *Rev. Mod. Phys.* **2018**, *90*, 035006. [CrossRef]
25. Pirandola, S.; Bardhan, B.R.; Gehring, T.; Weedbrook, C.; Lloyd, S. Advances in photonic quantum sensing. *Nat. Photonics* **2018**, *12*, 724–733. [CrossRef]
26. Tóth, G.; Vértesi, T. Quantum States with a Positive Partial Transpose are Useful for Metrology. *Phys. Rev. Lett.* **2018**, *120*, 020506. [CrossRef]
27. Polino, E.; Valeri, M.; Spagnolo, N.; Sciarrino, F. Photonic quantum metrology. *AVS Quantum Sci.* **2020**, *2*, 024703. [CrossRef]
28. Pál, K.F.; Tóth, G.; Bene, E.; Vértesi, T. Bound entangled singlet-like states for quantum metrology. *Phys. Rev. Res.* **2021**, *3*, 023101. [CrossRef]
29. Schleich, W. *Quantum Optics in Phase Space*; Wiley: Hoboken, NJ, USA, 2011. [CrossRef]
30. Weedbrook, C.; Pirandola, S.; García-Patrón, R.; Cerf, N.J.; Ralph, T.C.; Shapiro, J.H.; Lloyd, S. Gaussian quantum information. *Rev. Mod. Phys.* **2012**, *84*, 621–669. [CrossRef]

31. Adesso, G.; Ragy, S.; Lee, A.R. Continuous Variable Quantum Information: Gaussian States and Beyond. *Open Syst. Inf. Dyn.* **2014**, *21*, 1440001. [CrossRef]
32. Lvovsky, A.I., Squeezed Light. In *Photonics*; John Wiley and Sons, Ltd.: Hoboken, NJ, USA, 2015; Chapter 5, pp. 121–163. [CrossRef]
33. Maccone, L.; Riccardi, A. Squeezing metrology: A unified framework. *Quantum* **2020**, *4*, 292. [CrossRef]
34. Gatto, D.; Facchi, P.; Narducci, F.A.; Tamma, V. Distributed quantum metrology with a single squeezed-vacuum source. *Phys. Rev. Res.* **2019**, *1*, 032024. [CrossRef]
35. Gramegna, G.; Triggiani, D.; Facchi, P.; Narducci, F.A.; Tamma, V. Heisenberg scaling precision in multi-mode distributed quantum metrology. *New J. Phys.* **2021**, *23*, 053002. [CrossRef]
36. Gramegna, G.; Triggiani, D.; Facchi, P.; Narducci, F.A.; Tamma, V. Typicality of Heisenberg scaling precision in multimode quantum metrology. *Phys. Rev. Res.* **2021**, *3*, 013152. [CrossRef]
37. Triggiani, D.; Facchi, P.; Tamma, V. Heisenberg scaling precision in the estimation of functions of parameters in linear optical networks. *Phys. Rev. A* **2021**, *104*, 062603. [CrossRef]
38. Triggiani, D.; Facchi, P.; Tamma, V. Non-adaptive Heisenberg-limited metrology with multi-channel homodyne measurements. *Eur. Phys. J. Plus* **2022**, *137*, 125. [CrossRef]
39. Triggiani, D.; Tamma, V. Estimation with Heisenberg-Scaling Sensitivity of a Single Parameter Distributed in an Arbitrary Linear Optical Network. *Sensors* **2022**, *22*, 2657. [CrossRef]
40. Gatto, D.; Facchi, P.; Tamma, V. Heisenberg-limited estimation robust to photon losses in a Mach-Zehnder network with squeezed light. *Phys. Rev. A* **2022**, *105*, 012607. [CrossRef]
41. Triggiani, D.; Tamma, V. Estimation of the average of arbitrary unknown phase delays with Heisenberg-scaling precision. In *Proceedings of the Optical and Quantum Sensing and Precision Metrology II*; Scheuer, J., Shahriar, S.M., Eds.; International Society for Optics and Photonics, SPIE: San Francisco, CA, USA, 2022; Volume 12016, pp. 97–101. [CrossRef]
42. Proctor, T.J.; Knott, P.A.; Dunningham, J.A. Multiparameter Estimation in Networked Quantum Sensors. *Phys. Rev. Lett.* **2018**, *120*, 080501. [CrossRef]
43. Zhuang, Q.; Zhang, Z.; Shapiro, J.H. Distributed quantum sensing using continuous-variable multipartite entanglement. *Phys. Rev. A* **2018**, *97*, 032329. [CrossRef]
44. Matsubara, T.; Facchi, P.; Giovannetti, V.; Yuasa, K. Optimal Gaussian metrology for generic multimode interferometric circuit. *New J. Phys.* **2019**, *21*, 033014. [CrossRef]
45. Qian, K.; Eldredge, Z.; Ge, W.; Pagano, G.; Monroe, C.; Porto, J.V.; Gorshkov, A.V. Heisenberg-scaling measurement protocol for analytic functions with quantum sensor networks. *Phys. Rev. A* **2019**, *100*, 042304. [CrossRef]
46. Guo, X.; Breum, C.R.; Borregaard, J.; Izumi, S.; Larsen, M.V.; Gehring, T.; Christandl, M.; Neergaard-Nielsen, J.S.; Andersen, U.L. Distributed quantum sensing in a continuous-variable entangled network. *Nat. Phys.* **2020**, *16*, 281–284. [CrossRef]
47. Oh, C.; Lee, C.; Lie, S.H.; Jeong, H. Optimal distributed quantum sensing using Gaussian states. *Phys. Rev. Res.* **2020**, *2*, 023030. [CrossRef]
48. Grace, M.R.; Gagatsos, C.N.; Guha, S. Entanglement-enhanced estimation of a parameter embedded in multiple phases. *Phys. Rev. Res.* **2021**, *3*, 033114. [CrossRef]
49. Armen, M.A.; Au, J.K.; Stockton, J.K.; Doherty, A.C.; Mabuchi, H. Adaptive Homodyne Measurement of Optical Phase. *Phys. Rev. Lett.* **2002**, *89*, 133602. [CrossRef] [PubMed]
50. Monras, A. Optimal phase measurements with pure Gaussian states. *Phys. Rev. A* **2006**, *73*, 033821. [CrossRef]
51. Aspachs, M.; Calsamiglia, J.; Muñoz Tapia, R.; Bagan, E. Phase estimation for thermal Gaussian states. *Phys. Rev. A* **2009**, *79*, 033834. [CrossRef]
52. Berni, A.A.; Gehring, T.; Nielsen, B.M.; Händchen, V.; Paris, M.G.A.; Andersen, U.L. Ab initio quantum-enhanced optical phase estimation using real-time feedback control. *Nat. Photonics* **2015**, *9*, 577–581. [CrossRef]
53. Grace, M.R.; Gagatsos, C.N.; Zhuang, Q.; Guha, S. Quantum-Enhanced Fiber-Optic Gyroscopes Using Quadrature Squeezing and Continuous-Variable Entanglement. *Phys. Rev. Appl.* **2020**, *14*, 034065. [CrossRef]
54. Cramér, H. *Mathematical Methods of Statistics (PMS-9)*; Princeton University Press: Princeton, NJ, USA, 1946. [CrossRef]
55. Rohatgi, V.K.; Saleh, A.M.E. *An Introduction to Probability and Statistics*; John Wiley and Sons: Hoboken, NJ, USA, 2015.
56. Haar, A. Der Massbegriff in der Theorie der Kontinuierlichen Gruppen. *Ann. Math.* **1933**, *34*, 147–169. [CrossRef]
57. Nichols, R.; Liuzzo-Scorpo, P.; Knott, P.A.; Adesso, G. Multiparameter Gaussian quantum metrology. *Phys. Rev. A* **2018**, *98*, 012114. [CrossRef]
58. Demkowicz-Dobrzanski, R.; Górecki, W.; Guţă, M. Multi-parameter estimation beyond quantum Fisher information. *J. Phys. A Math. Theor.* **2020**, *53*, 363001. [CrossRef]
59. Stoica, P.; Marzetta, T.L. Parameter estimation problems with singular information matrices. *IEEE Trans. Signal Process.* **2001**, *49*, 87–90. [CrossRef]
60. Gross, J.A.; Caves, C.M. One from Many: Estimating a Function of Many Parameters; *J. Phys. A: Math. Theor.* **2021**, *54*, 014001 . [CrossRef]
61. Hiai, F.; Petz, D. *The Semicircle Law, Free Random Variables and Entropy*; Number 77; American Mathematical Soc.: Providence, RI, USA, 2000. [CrossRef]

62. Puchała, Z.; Miszczak, J.A. Symbolic integration with respect to the Haar measure on the unitary groups. *Bull. Pol. Acad. Sci. Tech. Sci.* **2017**, *65*, 21–27. [CrossRef]
63. Facchi, P.; Garnero, G. Quantum thermodynamics and canonical typicality. *Int. J. Geom. Methods Mod. Phys.* **2017**, *14*, 1740001. [CrossRef]
64. Popescu, S.; Short, A.J.; Winter, A. Entanglement and the foundations of statistical mechanics. *Nat. Phys.* **2006**, *2*, 754. [CrossRef]